

## **The NCEP Climate Forecast System Reanalysis**

Suranjana Saha<sup>1</sup>, Shrinivas Moorthi<sup>1</sup>, Hua-Lu Pan,<sup>1</sup>  
Xingren Wu<sup>3</sup>, Jiande Wang<sup>3</sup>, Sudhir Nadiga<sup>3</sup>, Patrick Tripp<sup>7</sup>,  
Robert Kistler<sup>3</sup>, John Woollen<sup>3</sup>, David Behringer<sup>1</sup>,  
Haixia Liu<sup>3</sup>, Diane Stokes<sup>1</sup>, Robert Grumbine<sup>1</sup>, George Gayno<sup>3</sup>, Jun Wang<sup>3</sup>,  
Yu-Tai Hou<sup>1</sup>, Hui-ya Chuang<sup>1</sup>, Hann-Ming H. Juang<sup>1</sup>, Joe Sela<sup>1</sup>, Mark Iredell<sup>1</sup>,  
Russ Treadon<sup>1</sup>, Daryl Kleist<sup>3</sup>, Paul Van Delst<sup>3</sup>, Dennis Keyser<sup>1</sup>, John Derber<sup>1</sup>,  
Michael Ek<sup>1</sup>, Jesse Meng<sup>3</sup>, Helin Wei<sup>3</sup>, Rongqian Yang<sup>3</sup>, Stephen Lord<sup>1</sup>.

Huug van den Dool<sup>2</sup>, Arun Kumar<sup>2</sup>, Wanqiu Wang<sup>2</sup>, Craig Long<sup>2</sup>, Muthuvel Chelliah<sup>2</sup>,  
Yan Xue<sup>2</sup>, Boyin Huang<sup>9</sup>, Jae-Kyung Schemm<sup>2</sup>, Wesley Ebisuzaki<sup>2</sup>, Roger Lin<sup>9</sup>,  
Pingping Xie<sup>2</sup>, Mingyue Chen<sup>2</sup>, Shuntai Zhou<sup>9</sup>, Wayne Higgins<sup>2</sup>

Cheng-Zhi Zou<sup>4</sup>, Quanhua Liu<sup>5</sup>, Yong Chen<sup>4</sup>, Yong Han<sup>4</sup>, Lidia Cucurull<sup>4</sup>,  
Richard W Reynolds<sup>8</sup>, Glenn Rutledge<sup>6</sup>, Mitch Goldberg<sup>4</sup>

Submitted to the Bulletin of the American Meteorological Society

January 12, 2010

<sup>1</sup> Environmental Modeling Center, NCEP/NWS/NOAA

<sup>2</sup> Climate Prediction Center, NCEP/NWS/NOAA

<sup>3</sup> Science Applications International Corporation (SAIC)

<sup>4</sup> Center for Satellite Applications and Research, NESDIS/NOAA

<sup>5</sup> Joint Center for Satellite Data Assimilation, NASA/NOAA

<sup>6</sup> National Climatic Data Center, NESDIS/NOAA

<sup>7</sup> Plurality Engineering and Scientific Solutions, LLC

<sup>8</sup> Cooperative Institute for Climate & Satellites, NOAA & North Carolina State University

<sup>9</sup> Wyle Information System

Corresponding Author: Dr. Suranjana Saha,  
NOAA Science Center  
5200 Auth Road, Camp Springs, MD 20746  
[Suranjana.Saha@noaa.gov](mailto:Suranjana.Saha@noaa.gov)

## **ABSTRACT**

The NCEP Climate Forecast System Reanalysis (CFSR) was completed over the 31-year period of 1979 to 2009 in January 2010. The CFSR was designed and executed as a global, high resolution, coupled atmosphere-ocean-land surface-sea ice system to provide the best estimate of the state of these coupled domains over this period. The current CFSR will be extended as an operational, real time product into the future.

The CFSR relative to most, if not all, previous reanalyses include (1) coupling of atmosphere and ocean during the generation of the 6 hour guess field, (2) an interactive sea-ice model, and (3) assimilation of satellite radiances by the Grid-point Statistical Interpolation scheme over the entire period. The CFSR global atmosphere resolution is ~38 km (T382) with 64 levels extending from the surface to 0.26 hPa. The global ocean is  $0.25^\circ$  at the equator, extending to a global  $0.5^\circ$  beyond the tropics, with 40 levels to a depth of 4737m. The global land surface model has 4 soil levels and the global sea ice model has 3 levels. The CFSR atmospheric model contains observed variations in carbon dioxide ( $\text{CO}_2$ ) over the 1979-2009 period, together with changes in aerosols and other trace gases and solar variations. With these variable parameters, the analyzed state will include estimates of changes in the Earth system climate due to these factors.

All available conventional and satellite observations were included in the CFSR.

Satellite observations were used in radiance form and were bias corrected with “spin up” runs at full resolution, taking into account variable  $\text{CO}_2$  concentrations. This procedure enabled smooth transitions of the climate record due to evolutionary changes in the satellite observing system.

CFSR atmospheric, oceanic and land surface output products are available at an hourly time resolution and  $0.5^\circ$  horizontal resolution. This reanalysis will serve many purposes, including providing the basis for most of NCEP Climate Prediction Center’s operational climate products by defining the mean states of the atmosphere, ocean, land surface and sea ice over the next 30-year climate normal (1981-2010) ; provide initial conditions for

historical forecasts required to calibrate operational NCEP climate forecasts (from week 2 to 9 months); and provide estimates and diagnoses of the earth's climate state, over the satellite data period, for community climate research.

Preliminary analysis of the CFSR output indicates a product far superior in most respects to the reanalysis of the mid-1990s. The previous NCEP reanalyses have been one of the most used NCEP products in history; there is every reason to believe the CFSR will supersede these older products both in scope and quality, because it is higher in time and space resolution, covers the atmosphere, ocean, sea ice and land, and was executed in a coupled mode with a more modern data assimilation system and forecast model.

## 1. *Introduction*

The first reanalysis at NCEP conducted in the 1990s, resulting in the NCEP/NCAR Reanalysis (Kalnay et al. 1996) or R1 for brevity, would ultimately cover many years, from 1948 to the present (Kistler et al. 2001). It is still being executed at NCEP, to the benefit of countless users for monthly, even daily, updates of the current state of the atmosphere. At the same time, other reanalyses were being conducted, namely, ERA15 (Gibson et al. 1997) was executed for a more limited period (1979-1993) at the European Center for Medium Range Forecasting (ECMWF), the Center for Ocean-Land-Atmosphere Studies (COLA) conducted a short reanalysis covering the period May 1982-Nov 1983 (Paolino et al. 1995) and the Goddard Space Flight Center (GSFC/NASA) conducted a reanalysis covering the period 1980-1994 (Schubert et al. 1997). ECMWF has conducted two subsequent reanalyses, ERA40 (Uppala et al. 2005) and ERA-Interim. The Japan Meteorological Agency (JMA) conducted JRA25 (Onogi et al. 2005), while GSFC is conducting MERRA (Bosilovich et al. 2008). A 100-year reanalysis using only surface pressure data is being conducted at ESRL/NOAA (Compo et al. 2006). The general purpose of conducting reanalyses is to produce multi-year global state-of-the-art gridded representations of atmospheric states, generated by a constant model and a constant data assimilation system. To use the same model and data assimilation over a very long period was the great advance during the 1990s, since gridded data sets available before 1995 had been created in real time by ever changing models and analysis methods, even by hand analyses prior to about 1965. The hope was that a reanalysis, made after real time, would help in climate studies by eliminating fictitious trends caused by model and data assimilation changes in real time. Although ‘climate’ was the consideration of overriding importance in justifying the enormous effort involved in conducting a multi-year reanalysis, the techniques used were those of numerical weather prediction (NWP). This climate–weather linkage lives on to this day. In particular, the analysis at any given time (t) is the result of a short forecast (the guess field), initialized

from a previous analysis (valid at time  $t-\Delta t$ ), amended by assimilating new observations available in a narrow window centered at  $t$ . The increment  $\Delta t$  is by convention 6 hours, but it could, in principle, be shorter or longer. Not surprisingly, most of the later reanalyses have been executed at operational weather prediction centers, the only exception being MERRA (Bosilovich et al 2008) conducted at GFSC/NASA.

One very time consuming, but rewarding, part of the first reanalyses was to assemble and convert older observational data sets into modern formats (such as BUFR), so they could be ingested into a modern modeling/data assimilation system. The data mining and preparation of these data sets is a lasting legacy, and they have been exchanged among national and international partners, and used in several more reanalyses (Uppala et al 2005; Onogi et al 2005, Bosilovich et al 2008). Every new reanalysis benefits from all previous reanalyses, and it is likely that mistakes are discovered and corrected to the benefit of the next user. This truly has been an ongoing activity in the international arena. Errors, big and small, were made in R1. Some were discovered and documented in:

<http://www.cpc.noaa.gov/products/wesley/reanalysis.html#problem>.

Kanamitsu et al (2002) executed a corrected version of R1, often called R2, covering only the satellite era of 1979-present. Importantly, in spite of correcting many errors and updating some components of the system, “only minor differences are found between R1 and R2 in the primary analysis variables, such as free atmospheric geopotential height and winds in the Northern Hemisphere extratropics” (Kanamitsu et al, 2002). Both R1 and R2 are being continued at NCEP, an activity sometimes referred to as the Climate Data Assimilation Systems, specifically CDAS1 for R1 and CDAS2 for R2. Running a CDAS implies certain technologies and computer algorithms are frozen in time, which in the face of ever changing data ingest and computer configurations can be a challenge, and ultimately would be impossible. A new reanalysis (ERA-Interim) executed presently at the ECMWF will be kept up to date (which ERA15 and its successor, ERA40, did not), thereby increasing its utility for real-time applications.

In this paper we only discuss *global* reanalyses. There has been one regional reanalysis conducted at NCEP, namely the North American Regional Reanalysis (NARR), see Mesinger et al. 2006, but this type of activity is outside the scope of this paper. However, it is worth noting that the CFSR has a global horizontal resolution of ~38km that nearly matches the 32km resolution of NARR over the limited area of the US.

It has become increasingly clear over the last two decades that a single reanalysis will not suffice. Not only is the data ingest subject to continual improvements, but the models used to create the guess field improve continually, and so do the data assimilation systems. Bengtsson et al (2007) advocate the idea of a permanent, or ongoing, analysis of the Earth system. These efforts would not only utilize state-of-the-art forecast and data assimilation methods, but in some configurations, the observations may even be reduced to whatever was available continuously over a century. Such kinds of reanalyses would be homogeneous over time (Compo et al. 2006) and could be used for climate studies at longer time scales.

New data assimilation techniques have been introduced since the mid 1990s, ranging from three dimensional variational (3DVAR), to four dimensional variational (4DVAR), and/or ensembles of analyses (EnKF) that produce not only an ensemble mean analysis, but also a measure of the uncertainty (spread), see Compo et al(2006) and Anderson et al(2009).

Although reanalyses may primarily be intended for climate studies, they also generate initial states that can be used to start integrations by a numerical weather prediction model, either for weather or climate, in real time or from some historical state. The desire to generate hindcasts (sometimes called Reforecasts or Retrospective forecasts) to calibrate subsequent real time forecasts, generated by the same model, have been under considerable development since the 1990s, and the availability of a reanalysis is key for these efforts to succeed (Hamill et al. 2006). For instance, when NCEP operationally implemented its first global coupled seasonal forecast system, namely the Climate Forecast System (Saha et al 2006) in August 2004 (hereafter referred to as CFS v1), a

complete Reforecast dataset was created over the period 1981-present to calibrate the real time operational forecasts. The initial states for this Reforecast effort were taken from R2 for the atmosphere and land, and from the Global Ocean Data Assimilation System (GODAS, forced by fluxes from R2) for the ocean. This relatively quick effort would have been inconceivable without R2 being available.

Since the operational CFS v1 implementation in 2004 (Saha et al. 2006), CPC and many other users have utilized the four times daily seasonal integrations for their monthly and seasonal forecast products. However, the system has many internal inconsistencies. For instance, the R2 atmospheric initial states are made with technology from the 1990s, while the atmospheric model component of CFS is from a decade later. Thus the initial states and forecast model are inconsistent; a situation which leads to loss of skill during the early part of the integrations. One of the major goals of executing the new CFSR was to create initial states for the atmosphere, ocean, land and sea-ice which are both state-of-the-art, and as consistent as possible with the next version of the CFS v2 which is to be implemented operationally at NCEP in 2010. Given the pace of model and data assimilation development, such a reanalysis will be needed roughly every 5-10 years.

We have now completed the CFSR for the 31-year period 1979-2009. It took almost two years to accomplish this feat. The primary novelties of this latest reanalysis are a) coupling to the ocean during the generation of the 6 hour guess field, b) an interactive sea-ice model, and c) assimilation of satellite radiances for the entire period. In addition, the much higher horizontal and vertical resolution (T382L64) of the atmosphere, model and assimilation improvements over the last 10-15 years, and the use of prescribed CO<sub>2</sub> concentrations as a function of time, should make for substantial improvements over R1 and R2 (which were at T62L28 resolution). Another major advance was the real time monitoring that took place during the execution of the CFSR. Thousands of graphical plots were generated automatically at the end of each reanalyzed month, and displayed on the CFSR website in real time. Many scientists from both CPC and EMC monitored different aspects of the reanalysis during this 2 year process. There were many times that

the reanalysis was halted to address concerns that something may have gone wrong, and many corrections and restarts were made to ensure that the process was done correctly. This extremely large “atlas” of plots depicting nearly all aspects of the CFSR will be open to the public at: <http://cfs.ncep.noaa.gov/cfsreanl>

It is worth noting that the analysis system used in CFSR for the atmosphere, the Gridpoint Statistical Interpolation (GSI) scheme, is nearly the same as the one used by MERRA at GSFC/NASA. The MERRA atmosphere-only reanalysis is being conducted over the same years with nearly the same input data.

The lay-out of the paper is as follows. There are sections devoted to (2) the observations, (3) the atmosphere, (4) the ocean, (5) the sea-ice, (6) the coupler, (7) the land and (8) the execution. Sections 9 and 10 deal with preliminary analysis of the CFSR, that resulted from the ongoing monitoring. Far more substantial analyses of this reanalysis will be forthcoming in a dozen or so journal articles that are currently under consideration. In the concluding section 11, we look ahead to the completion of the CFSR Project, namely the Reforecasts (the second R in CFSRR) over the 1980-present period that will be initialized by the CFSR. This project is currently in progress. We will also allude to the possibility of going back to 1948, or further, with a reduced resolution of the CFSR system, such that both R1 and R2 have a modern successor, and the old technologies can be retired gracefully. Since some of the sections are highly technical and detailed, the reader is referred to Appendix A for acronyms. Appendix B describes some of the data that is available for distribution, while Appendix C gives information about data access. It must be noted here that data from this reanalysis has been archived at unprecedented spatial ( $0.5^\circ \times 0.5^\circ$ ) and temporal (hourly) resolution, exceeding over 250 Terabytes.

## **2. *The Observations***

Reanalysis projects depend upon both historical and operational archives of observations, and newly reprocessed sets of observations being produced at meteorological research centers around the world. Historical observations are often found stored in obsolete formats and mediums, at various research institutions and national archives in various



stages of incompleteness and/or disarray. Fortunately, there is a wealth of modernized historical observation datasets currently available, which have been produced for a series of major global reanalysis projects, as outlined in the Introduction. A great deal of historical data mining and archeology has been accomplished, particularly at NCAR, NCDC and NESDIS, in preparation for these and other climate research projects (Jenne and Woollen, 1994, Woollen and Zhu, 1997). In addition, a number of important modern datasets have been reprocessed and improved from operational archives at international centers, such as ECMWF, NCEP, ESA, EUMETSAT, JMA, and BOM.

## **2.1    *Observation Sources***

The CFSR project will be completed over two epochs of meteorological observing periods. Initially, CFSR has analyzed the “TOVS” or “modern” era, from 1978 through the present. Subsequently the CFSR will be extended back in time to 1947 or further. Most of the observations for the period 1948 through 1978 are available as products of the previous reanalysis preparation efforts at NCEP. Observations assimilated from 1978 through 1997 were drawn from sources which were either copied from previous projects, or from datasets improved or made available since previous NCEP reanalysis projects were completed. The observations for the most recent period of the CFSR, 1997-2009, have mostly been drawn directly from the NCEP operational run history archives, which have been saved in tape silos beginning when the NCEP modernized real-time BUFR database was implemented.

## **2.2    *Observation Preparation***

The bulk of the work involved in data preparation for CFSR was invested in 1) data movement and archiving of large volume data sets, especially those from satellite based observing systems, and 2) the combination of datasets containing overlapping contents. Once a database of all observations to be included in the CFSR input were assembled in the correct BUFR formats, and with duplicate observations resolved, current NCEP operational data prep codes were used to create online data dump directories from which the reanalysis processing scripts would access the input datasets for assimilation. The

online archive of data for reanalysis created for CFSR processing was convenient because it allowed corrections or additions to be made to the dumped data files quickly and cleanly. This proved to be very useful in several instances where significant problems with certain data types were able to be resolved very quickly and without impacting the processing schedule. An annotated diagram of the distribution of the 8 terabytes of data found in the 30 year online data dump archive is shown in Figure 1.

### **2.3     *Conventional Observing Systems in the CFSR***

The sections in part 2.3 briefly describe some specific details about the conventional observing systems used in the CFSR, their sources, characteristics, and preparation. Each section also references one or more charts which illustrate the observing system performances of the data described, with respect to the quality control reactions and the monthly RMS and mean fits to the analysis and first guess backgrounds over the 31 year period, 1979 through 2009.

#### **2.3.1   *Radiosondes and Pibals***

Many sources of radiosonde and pilot balloon data have been made available for reanalysis projects. From 1948 through 1997 a number of archives were combined for the CFSR assimilation, including operational archives from NMC/NCEP, ECMWF, JMA, USAF, NAVY, along with various other military, academic, and national archives collected at NCAR and NCDC over the years. The main issue in combining radiosonde archives is resolving overlapping, or duplicate, soundings. For the CFSR radiosonde preparation, duplicates were resolved by merging the contents of duplicate soundings instead of picking one sounding from one of the sources and discarding the others, as has been commonly done in reanalysis projects to date. This merging process was completed for the entire set of radiosonde data for the years 1948 through 1997. Starting in May 1997, the NECP operational run history archive supplied the bulk of CFSR radiosonde data. Figures 2-3 show the performance of 500mb radiosonde temperatures and winds, respectively, in the CFSR.

#### ***2.3.1.1 African Monsoon Multidisciplinary Analysis (AMMA) Special Observations in CFSR***

A special observation program known as AMMA has been ongoing since 2001, focused on reactivating, renovating, and installing radiosonde sites in West Africa. Between 2004 and 2006 the number of radiosonde sites in that region jumped from 5 to 21 active stations. During the summer of 2006 an intensive observing period was carried out in which some 7000 soundings were made, representing the greatest density of radiosondes ever launched in the region (Kadi, 2009) . In addition these soundings were augmented by high resolution profiler and dropsonde observations. The CFSR was able to include much of this special data in 2006, thanks to an arrangement with the ECMWF and the AMMA project.

#### ***2.3.2 Aircraft and ACARS data***

Aircraft have been used for many years to make observations of winds and temperatures along their flight paths. The bulk of CFSR aircraft observations are taken from the US operational NWS archives which start in 1962, and are continuous through the present time. A number of other archives from military and national sources have been obtained and used to provide data which is not represented in the NWS archive. Very useful datasets have been supplied, in particular, by NCAR, ECMWF, and JMA. The automatic ACARS aircraft observations enter the CFSR in 1992. Figures 4-5 show the performance of aircraft wind, and ACARS wind, respectively.

#### ***2.3.3 Surface Observations***

The US NWS operational archive of ON124 surface synoptic observations is used beginning in 1976 to supply land surface data for CFSR. Prior to 1976, a number of military and national archives were combined to provide the land surface pressure data for the CFSR. All the observed marine data from 1948 through 1997 have been supplied by the COADS datasets. These data include ship reports, buoys, and other marine observing platforms. Starting in May 1997 all surface observations are taken from the NCEP operational archives. METAR automated reports also start in 1997. Very high

density MESONET data is included in the CFSR database starting in 2002. Although these observations are not assimilated, they are available in CFSR files for validation. Figures 6-9 show the performance of SYNOP and METAR surface pressure, and marine temperature and wind, respectively.

#### **2.3.4 PAOBS**

PAOBS are bogus observations of sea level pressure created at the Australian BOM from the 1972 through the present. They were initially created for NWP to mitigate the extreme lack of observations over the Southern Hemisphere oceans. Studies of the impact of PAOB data (Seaman and Hart, 2003) show positive impacts on SH analyses, at least until 1998 when ATOVS became available. CFSR assimilated PAOBS through 1998, although there are indications in the performance chart (Figure 10) that PAOBS would be useful until 2002. CFSR had GTS PAOBS until 1985, included reprocessed PAOBS from JRA-25 datasets, 1985-2000, then PAOBS from NCEP operational archives through the present.

#### **2.3.5 Satellite Wind (SATOB) Observations**

Atmospheric motion vectors derived from geo-stationary satellite imagery are assimilated in the CFSR beginning in 1979. The imagery from GOES, METEOSAT, and GMS satellites provide the observations used in CFSR, mostly obtained from US NWS archives of GTS data. GTS archives from JMA were used to augment the NWS set through 1993 in the R1. Reprocessed high resolution GMS SATOB data were specially produced by JMA for the JRA-25 reanalysis, covering the years 1987 through 2003. These were made available for the CFSR project, and replaced the GTS GMS data in the assimilation database for that period. MODIS polar wind data is obtained from the NCEP archives for the CFSR starting in late 2004. Recently, a 27 year dataset of northern and southern polar winds, derived from AVHRR on TOVS, ATOVS, and METOP satellites, has been created for use in reanalysis projects (Dworak and Kay, 2007), although not in time for CFSR preparation. Figures 11-13 show the performance of METEOSAT, GMS, and GOES satellite wind observations, respectively.

### **2.3.6 SSM/I Ocean Surface Wind Speed**

Microwave imager data from the SSM/I instrument is processed to derive ocean surface wind speed observations. SSM/I microwave radiance datasets, from DMSP satellites, were obtained from NCDC starting in 1993. Beginning in 1997 the DMSP data is available in the NCEP archives. The SSM/I brightness temperature data were converted to wind speeds by a neural net algorithm developed at NCEP and used in NCEP operations (Krasnopolsky et al, 1996, Gemmill and Krasnopolsky, 1999, Yu et al, 1997).

### **2.3.7 Scatterometer Winds**

Ocean surface wind datasets have been available from the European Space Agency ERS-1/AMI scatterometer since 1991, and the ERS-2/AMI instrument since 1996. ESA has recently reprocessed ocean surface wind vectors from the ERS-1 and ERS-2 satellite archives, covering the years 1991 through 2007. The reprocessed data was obtained for the entire period, and assimilated in the CFSR. Figure 14 shows the performance of ERS/AMI ocean surface wind data in the CFSR. The NASA QuikSCAT Seawinds scatterometer was launched in 1999. Seawinds ocean surface vector wind data from the NCEP operational archives was assimilated in CFSR from 2001 until it went non-operational in 2009. The NRL WINDSAT scatterometer data is assimilated in CFSR starting September 2008.

## **2.4 Satellite Radiance Based Observing Systems in the CFSR**

The NCEP operational GDAS has directly assimilated satellite radiances for a number of years, but CFSR is the first NCEP global reanalysis to do so. The historical TOVS and ATOVS archives were obtained from the NESDIS web based CLASS online archive. All the other radiance data were obtained from the NCEP operational archives. Figure 15 is a CFSR usage chart of radiance measuring instruments illustrating the time period during which each instrument was assimilated.

### **2.4.1 TOVS radiances**

The CFSR assimilated radiance data from satellite sounders with TOVS instruments on board nine NOAA polar-orbiting satellites from TIROS-N to NOAA-14 starting in 1978.

The 1B datasets were calibrated using operational calibration coefficients stored in the files to convert the raw data counts into brightness temperatures for the assimilation. The TOVS radiance data was assimilated in the CFSR through 1998, although the TOVS data remains in the CRSR files until 2007. Figures 16-18 show the performance of HIRS/2 channel 12, MSU channel 3, and SSU channel 1, for the full period they are in the CFSR data dumps.

#### **2.4.2 *Recalibrated MSU Radiances***

Currently, there are two sets of calibration processing codes that are optional for generating MSU level-1c radiances from the level-1b files. The first one is known as the ‘NOAA operational calibration’, in which calibration coefficients used for converting raw counts data to level-1c radiances were obtained from pre-launch chamber test datasets (Mo et al. 2001). It is well known that the NOAA TOVS satellites suffered through an orbital-drift problem, which caused an onboard warm target temperature to have a large variability owing to differences in sun heating on the instrument (Zou et al. 2006). The warm target was designed for calibrating the MSU raw observations, however, the NOAA pre-launch-based operational calibration did not account for the large warm target temperature variability. As a consequence, a part of this variability remained in the NOAA operational calibrated level-1c radiances. This is the so-called “warm target signal contamination” in the level-1c MSU radiances. NESDIS developed a post-launch calibration algorithm using simultaneous nadir overpasses (Zou et al., 2006, 2009). The new calibration algorithm largely removed inter-satellite biases and nonlinear bias drift caused by the warm target effect resulting in more accurate radiances. By the time the CFSR was begun, recalibration of MSU channels 2, 3, 4 from NOAA-10 to NOAA-14 was completed, and the recalibrated data were included in CFSR. Figure 19 presents MSU global average first guess brightness temperatures for MSU channels 1-4, for the entire period they were assimilated. The effect of recalibration of channels 2-4 on NOAA-10 to NOAA-14 is clearly seen.

#### **2.4.3 *ATOVs radiances***

Advanced TOVS instrumentation became available in 1998 with the launch of the NOAA-15 satellite and continued on NOAA-16 and NOAA-17. Level 1b ATOVS datasets for HIRS/3, AMSU-a, and AMSU-b, were obtained from NESDIS/CLASS through June 2002, and calibrated with operational coefficients for use in CFSR. Starting in July 2002 calibrated ATOVS data was supplied from the NCEP operational archives. Evidence from a number of studies (e.g. English et al. 2000), and from operational experience, suggests that ATOVS data has been an important factor in large forecast improvements seen over the last decade. Figures 20-21 show the performance of HIRS/3 channel 5, and AMSU-a channel 5 in the CFSR.

#### **2.4.4 *GOES Radiances***

GOES sounder radiances have been available since the launch of GOES-8 in 1994. CFSR started assimilating superobed sounder radiances from GOES-8 and GOES-9 in 1997 when the NCEP operational archive began. Full resolution sounder data was introduced into CFSR assimilation in 2007.

#### **2.4.5 *AQUA - AIRS AMSU-a, and AMSRE data***

The AQUA satellite was launched by NASA in May 2002. Several advanced infrared sounders from NASA were on board, AIRS and AMSU-a, along with a microwave scanning radiometer, AMSR-E, built by the Japanese NSDA. The AIRS instrument is the first hyper-spectral sounder in space, which measures bandwidth brightness in approximately 3000 wave bands. For operational use, the AIRS data was thinned by a factor of 40 in the horizontal arrangement, and by a factor of 10 in the bandwidth measurements. CFSR has followed this protocol and assimilated the AIRS data in this configuration since late 2004. In mid-2007 the operational NCEP GDAS considered all 9 AIRS spots in each AMSU field of view for assimilation and the CFSR followed suit. Figure 22 shows the AIRS channel 215 (channel 92 in the 281 channel subset) performance in the CFSR.

#### **2.4.6 *METOP - IASI, AMSU-a, and MHS data***

MetOp is Europe's first polar-orbiting satellite dedicated to operational meteorology. It represents the European contribution to a new cooperative venture with the United States providing data that will be used to monitor climate and improve weather forecasting. The MetOp-A was launched in October 2006 and became operational in May 2007. One of the instruments carried onboard MetOp-A is IASI, the highest resolution infrared sounding interferometer currently in orbit. IASI observes the atmosphere in the infra-red (3.7 - 15.5  $\mu\text{m}$ ) on 8461 channels. Also important for the CFSR is the AMSU-a sounder, and the MHS instrument which replaces the AMSU-b instrument. MetOp-A data was introduced into the CFSR assimilation in Jan 2009.

## **2.5     *CHAMP/COSMIC GPS Radio Occultation data***

One of the innovative aspects of the CFSR is the assimilation of Global Positioning System (GPS) Radio Occultation (RO) observations. The GPS RO technique utilizes the transmission from the GPS satellites to a Low Earth Orbit (LEO) satellite to actively sound the atmosphere. By detecting the “bending” of the GPS signal as it passes through the atmosphere (see eg. Kursinski et al., 1997, Rocken et al., 1997), some properties of the thermodynamic state of the atmospheric can be retrieved. The derived soundings have equal density and good accuracy over land and oceans in both hemispheres. In addition, soundings are provided at high vertical resolution ( $\sim 100$  m), which complements the use of the nadir satellite sounders. The CFSR assimilates GPS RO observations from the CHAMP mission (Wickert et al. 2001) from May 2001 to December 2007. Observations from the US-Taiwan COSMIC mission, launched in April 2006, are assimilated starting in July 2006. COSMIC is a six LEO microsatellite constellation which provides around 2,000 daily worldwide atmospheric soundings, uniformly distributed over the globe, which produce a positive impact in the NCEP operational GDAS ((Cucurull and Derber 2008, Cucurull 2009).

## **2.6     *Observation Quality control***



The CFSR uses the NCEP operational observation quality control procedures which are summarized in Table 1. For more specific information on NCEP QC procedures see also: [http://www.emc.ncep.noaa.gov/mmb/data\\_processing/prepbuf.fr.doc/document.htm](http://www.emc.ncep.noaa.gov/mmb/data_processing/prepbuf.fr.doc/document.htm) .

Performance charts including QC summaries for many all observing systems used in the CFSR can be found on the CFSR website <http://cfs.ncep.noaa.gov/cfsreanl> .

### **3. *The Atmosphere***

#### **3.1 *The Atmospheric Data Assimilation***

R1 and R2 were run with modified versions of the mid-1995 operational GDAS system, based on a sigma coordinate spectral prediction model and the 3D VAR SSI (Parrish and Derber, 1992; Derber, Parrish and Lord, 1991). These systems assimilated a fixed set of conventional observations along with TOVS/ATOVS temperature retrievals (Smith, et. al, 1979). Both projects continue as CDAS.

##### **3.1.1 *Evolution of the NCEP GDAS***

In anticipation of the eventual replacement of TOVS by the ATOVS instruments in 1998, retrieval assimilation was replaced with the direct assimilation of cloud cleared radiances in the SSI (Derber and Wu, 1998, McNally, et. al. 2000). Sources of radiance data other than ATOVS, such as GOES sounder, AIRS, and AMSR-e, were added to the SSI assimilation as they became available. The operational GDAS underwent another major upgrade in May 2007 involving both the method of data assimilation and the vertical coordinate of global prediction model. The sigma coordinate system (Phillips, 1957) of the global spectral model, in use since the early 1980's, was replaced with a hybrid sigma-pressure system (Juang, 2005). The SSI was replaced by the GSI (Kleist, et. al, 2009), which permits a more general treatments of the background errors such as flow dependence and anisotropic structure functions (Purser et al. 2003a, Purser et al. 2003b). The 2007 operational GSI would form the basis for the CFSR.

##### **3.1.2 *Use of the SSU in CFSR***

The SSU instruments, onboard the majority of TOVS satellites, provide unique 29-year observations for studying stratospheric temperatures. The SSU is a step-scanned infrared

spectrometer employing a selective absorption technique to make measurements at the top of the Earth's atmosphere in three channels at the 15 micron carbon dioxide ( $\text{CO}_2$ ) absorption band. The three SSU channels have the same frequency but different  $\text{CO}_2$  cell pressures. With the three modulated cell pressures, the weighting function of the original 15 micron carbon dioxide absorption band is shifted up and split into three weighting functions, approximately located at 15, 5, and 1.5 hPa, for SSU channels 1, 2, and 3 respectively.

However, historical use of the SSU radiances posed a challenge due to this complicated sensor response and a leaking problem in the instrument's  $\text{CO}_2$  cell pressure modulator that caused the radiances from each satellite to exhibit a unique drift in time (Kobashi, et. al, 2009). These issues, plus the inevitable obsolescence of the SSU instrument, resulted in its omission from the original and subsequent versions of the CRTM. A line-by-line transmittance model was applied to generate the sensor response function for a fast transmittance model, and the CRTM, with its advanced surface emissivity model and radiative solver (Liu and Weng, 2006) was used to quantitatively correct the leaking effect. The transmittance model in the CRTM uses a set of coefficients depending on the  $\text{CO}_2$  cell pressures to compute atmospheric transmittances. By comparison to the detailed line-by-line calculation, the root mean square error due to the fitting and interpolation of the  $\text{CO}_2$  cell pressure in the fast transmittance model is less than 0.1 K (Liu and Weng, 2009).

The SSU radiative transfer calculations were then compared to the SSU radiances from NOAA-14. The input temperature profiles are taken from the Earth Observing System (EOS) Aura microwave limb sounding (MLS) product for November 2004, a completely independent data source. The MLS observes thermal microwave emission from Earth's 'limb', which can be used to estimate atmospheric composition, temperature, humidity and cloud ice with a high vertical resolution. The MLS temperature product precision throughout the stratosphere is generally less than 1 K. More than 7000 match-up data points are found, and all the data points are analyzed. Figure 23 illustrates the results. The

SSU and the MLS measurements are very consistent. The bias and rms error in the brightness temperature calculations at the SSU channel 1 (left panel), 2 (middle panel), and 3 (right panel) are less than (or equal to) 1 degree K.

### **3.1.3 *Development of the GSI for the CFSR***

Several innovative features were built into the operational version of the GSI implemented in May 2007, but not initially activated. In configuring the CFSR, it was decided to anticipate the activation of three of these features and include them from the start. The first of these was to apply flow dependence to the background error variances, (Kleist, personal communications) in an effort to improve upon the climatological estimates previously in use (Wu et al. 2002, Kleist et al. 2009). The algorithm takes advantage of the fact that tendency information for the guess fields (3, 6 and 9 hour model forecasts) is readily available. The static variances undergo a simple rescaling based on the 6-hr tendency in the model forecast, where the variances are increased (decreased) where the model tendencies are relatively large (small). The rescaling is performed level by level for each variable independently, and done in such a way as to approximately preserve the global mean variance as specified by the static estimate (i.e. it is not designed to increase or decrease the global mean error variance on a cycle to cycle basis). This procedure transforms the simple latitude and height dependent fixed variances into a fully three-dimensional, time-varying estimate as shown in Figure 24. The second dormant GSI feature enabled in the CFSR is FOTO, or first order time interpolation to the observation (Rančić, et. al, 2008). Many observation types are available throughout 6 hour assimilation window, but 3DVAR does not account for the time aspect. FOTO is a step in this direction. The observational operators in the minimization procedure are generalized to use time tendencies of state variables. The extra calculation slows convergence, but results in improved fits to the observations for the cost of a few additional iterations. FOTO also includes a parameterization of the planetary boundary layer, which was viewed as being complimentary to the ocean-atmosphere coupling in the CFSR.

The third GSI feature enabled in the CFSR is non-linear variational quality control (VarQC), (Anderson and Järvinen, 1999), which replaces the OIQCBUFR program (Woollen, 1991) that was used in R1 and R2 (Woollen et al. 1994). In the VarQC procedure, conventional GSI observation innovations must first pass gross error checks. Then an innovation weight is computed based on its consistency with the solution of the variational minimization based on all available observations, including radiances, with additional input coming from the probabilities of error for the various observations. Any observation with a weight of .25 or greater is effectively used in the minimization. Typically in a pass/fail QC procedure, an observation with a comparable weight of less than .7, or so, would be rejected from the process completely.

Another innovative feature of the CFSR GSI is the use of the historical concentrations of carbon dioxide (CO<sub>2</sub>, <http://gaw.kishou.go.jp>) when the historical instruments were retrofit into the CRTM, which originally only included TOVS, MSU and HIRS2 for NOAA-11 and NOAA-14. Table 2 lists the values of CO<sub>2</sub> concentrations (ppmv) used in the CRTM to calibrate satellites back to TIROS-N in 1979.

#### ***3.1.4 Satellite bias correction spin up for CFSR***

The direct assimilation of radiances represents one of the major improvements of the CFSR over R2. However, substantial biases exist when observed radiances are compared to those simulated by the CRTM depiction of the guess. These biases are complicated and relate to instrument calibration, data processing and deficiencies in the radiative transfer model. A variational satellite bias correction scheme was introduced by Derber and Wu (1998) to address this issue when direct assimilation of radiances began at NCEP. This scheme has been continually developed and is used in the GSI system adapted for the CFSR. Before the radiances of a new instrument can be assimilated, its unique set of starting bias corrections must be determined by a separate spin-up assimilation. In the case of CFSR, each set of historical instruments in Figure 21 required an individual spin-up. Since the TOVS instruments had never been assimilated by a GSI based GDAS, a preliminary set of tests were run (not shown) which determined that a 3 month spin-up

was required prior to the introduction of those historical instruments in the CFSR. A detailed explanation of this important step is given in Appendix D.

Examples of the bias correction values actually applied to the CFSR over the TOVS period of the CFSR, 1979-1998, may be seen in global averaged, 4 times daily averaged time series for MSU channels 1-4 and SSU channels 1-3 in Figure 27. (The spin-up of the SSU channels was done at the same time). The one measure of the successful spin-up procedures is the lack of discontinuities in the transitions between successive instruments. The break in the MSU time series are a result of the recalibration that was applied beginning in 1986 as noted in section 2.4.2.

### **3.1.5 *Transition to Real Time CFSR***

Once CFSR stream 4 (see section 8 for execution layout) completed February 2009, a decision point was reached. The operational GSI had gone through several upgrades during the CFSR execution, the latest in February being a major addition to the CRTM to simulate the hyper-spectral channels of the IASI instrument, onboard the new ESA METOP satellite. The IASI radiances had become operational in March 2009.

In order to continue to meet the goal of providing the best available initial conditions to the CFS, in the absence of staff and resources to maintain the CFSR GSI into the future, it was decided to make the transition to the CDAS mode of CFSR. The operational GSI, present and future implementations, would replace the CFSR GSI, and the coupled prediction model would be “frozen” to that of the CFS v2. Historical observational datasets would be replaced with the operational data dumps. One consequence of the switch to the operational GSI would be, because the anticipated operational implementation of FOTO did not happen, the period of March 2009 forward would be run without FOTO.

### **3.1.6 *QBO problem in the GSI***

The QBO can only be fully depicted in assimilation systems by sufficient direct wind observations, since the underlying physical mechanism is based on the dissipation of upwardly propagating gravity waves (Lindzen and Holton, 1968) which filtered out by

the hydrostatic assumption. Soon after CFSR production began, it was noticed that streams 2 and 3 (see section 8 for execution layout), completely missed the QBO wind transition. This was unexpected based on the ability of R1, R2 and CFSR streams 1 and 4 (starting in 1979 and 2004 respectively) to capture the QBO wind patterns. While searching for a comprehensive solution, it was noted that the ERA-40 tropical stratospheric wind profiles were readily available for the streams in question, included the stratospheric layers needed, and, qualitatively, adequately depicted both the QBO and semi-annual oscillation. In order that the streams could proceed with a reasonable QBO signature it was decided to use the ERA-40 stratospheric wind profiles as bogus observations for the period from Jul 1, 1981 to Dec 31, 1998.

Streams 1 and 4 had a common denominator. Both had more vertically resolved and temporally complete tropical radiosonde wind observations at the beginning of the stream, than those in stream 2. Stream 1 benefited from the enhanced FGGE observation system, and stream 4 from the automation of modern radiosonde data collection which results in more reports reaching the GTS, and more stratospheric levels in the individual reports. The solution to this problem became apparent from consultations with several GMAO MERRA team members, after determining that the MERRA reanalysis, which uses the same GSI assimilation component, depicted the QBO very well. Prior to starting the MERRA project, GMAO had experienced a similar problem analyzing the QBO in an earlier grid point analysis system. The problem was resolved by enlarging the horizontal length scale of the zonal wind correlation function in the tropical stratosphere (Gaspari, et.al, 2006). When the GMAO assimilation system was switched to the GSI, the tropical stratospheric stream function variances of the background error reflected the changes made to fix the problem in the earlier system. When comparable background error variances were tested in the GSI for a case where the CFSR had failed to capture the QBO, the wind transition was successfully analyzed (not shown).

### **3.1.6 *Tropical Cyclone Processing***

As global prediction models have become increasingly more highly resolved and have included more realistic boundary layer, water cycle and radiative parameterizations, the genesis and evolution of tropical storms has dramatically improved. Equally important are the ability of improved observing systems and analysis procedures to depict the details of tropical storm circulations and their accompanying steering flow.

The first global reanalysis to assimilate historical tropical storm information was the JRA-25 reanalysis (Onogi, et.al. 2007). It assimilated synthetic wind profiles (Fiorino, 2002) surrounding the historical storm locations of Newman, 1999.

A unique feature of the CFSR is its approach to the analysis of historical tropical storm locations. The CFSR applied the NCEP tropical storm relocation package (Liu et. al., 1999), a key component of the operational GFS analysis and prediction of tropical storm. By relocating a tropical storm vortex to its observed location prior to the assimilation of storm circulation observations, distortion of the circulation by the mismatch of guess and observed locations is avoided. Fiorino (personal communication) provided the CFSR with the historical set of storm reports converted into the operational format provided to NCEP by the NOAA/NWS National Hurricane Center and the US Navy Joint Typhoon Warning Center. Figure 28 shows the yearly total of tropical storm reports stacked by the 8 geographical basins.

A measure of the ability of assimilation to depict observed tropical storms is to quantify whether or not a reported storm is detected in the assimilation guess. Figure 29 shows the time series of the percentage of detected tropical storms plotted globally and for selected Northern Hemisphere basins: the Atlantic-Caribbean (ATL), the western Pacific (W Pac), and Eastern Pacific (E Pac). A noticeable improvement starts in 2000 coincident with the full utilization of the ATOVS satellite instruments, such that between 90-95% of reported tropical storms are detected.

### **3.2    *The Atmospheric Forecast Model***

The first NCEP/NCAR Reanalysis (R1, Kalnay et al., 1996; Kistler et al., 2001) was based on the operational medium range forecast model of January 1995. R1 had a T62 horizontal

resolution with 28 sigma layers in the vertical with the top layer near 3 hPa. Since then the NCEP global forecast system (GFS) atmospheric model (AM) has undergone many major revisions and improvements. The current operational CFS (Saha et al., 2006) is based upon the GFS AM of 2003. Changes included in the operational CFS, after R1, are to the boundary layer vertical diffusion (Hong and Pan 1996), simplified Arakawa-Schubert cumulus convection (Pan and Wu, 1995; Hong and Pan 1998) with cumulus momentum mixing and orographic gravity wave drag (Kim and Arakawa 1995, Alpert et al. 1988, 1996). The land surface model used in the CFS is based on the two-layer Oregon State University (OSU) land model (Pan and Mahrt, 1987). In addition, the cloud condensate is a prognostic variable (Moorthi *et al.* 2001) with a simple cloud microphysics parameterization (Zhao and Carr 1997, Sundqvist et al., 1989). Both large-scale condensation and detrainment of cloud water from cumulus convection provide source for cloud condensate. The fractional cloud cover used in the radiation calculation is diagnostically determined by the predicted cloud condensate based on the approach of Xu and Randall (1996). The CFS also has ozone as a prognostic variable with a simple parameterization for ozone production and destruction based on ten day mean climatological data supplied by NASA/GSFC. The shortwave (SW) radiation is parameterized following the NASA approach (Chou et al., 1998, Hou et al., 1996, 2002) and the longwave (LW) radiation following GFDL (Fels and Schwarzkopf, 1975; Schwarzkopf and Fels, 1991). Both radiation parameterizations use random cloud overlap with shortwave being called every hour and longwave every three hours. This version of GFS AM produced realistic moisture prediction, which may have contributed to a better depiction of no-rain areas and much improved tropical and hurricane track prediction.

Subsequent improvements to the GFS AM include fast and accurate LW radiation parameterization based on Rapid Radiative Transfer Model (RRTM) developed at Atmospheric and Environmental Research Inc. (AER, Mlawer et al., 1997). It is also coupled to a four-layer Noah land-surface model (Ek et al., 2003). In addition to gravity-wave drag, the model now includes a parameterization of mountain blocking (Alpert, 2004)



following the subgrid-scale orographic drag parameterization by Lott and Miller (1997). A two-layer sea-ice model (Wu et al., 2005) is also part of the current operational GFS AM and the model now uses the Earth System Modeling Framework (ESMF) base modern computer algorithms (Collins et al., 2005). Update of the ozone production and destruction terms is done by using the monthly mean data provided by NRL (McCormack et al., 2006). The horizontal resolution is T382 with 64 hybrid vertical layers with the top layer  $\sim 0.2$  hPa. The vertical structure of the model levels for the 28 sigma layer model (left panels) used in R1 and the 64 layer sigma-pressure hybrid model (right panels) used in the GFS/CFSR is shown in Figure 30 as a cross section at 90E, and as a cross section across North America at 40N in Figure 31. The top panels in both figures are plotted as a function of pressure to emphasize the resolution in the troposphere. In the bottom panels, the vertical axis is  $\log(\text{pressure})$  to emphasize the stratosphere. It can be seen that the new vertical structure adds more resolution in the stratosphere and does not have the topography signature above 100 hPa.

The version of GFS AM used in the CFSR has additional changes. It now uses enthalpy ( $C_pT$ ) as a prognostic variable in place of virtual temperature (Akmaev and Juang, 2008, Juang 2009) with a generalized hybrid vertical coordinate with sigma-pressure option. The vertical discretization is based on the generalized hybrid vertical coordinate (Juang, 2005). It also uses RRTM shortwave radiation with maximum random cloud overlap (Iacono et al., 2000; Clough et al., 2005) and a simple modification of shallow convection/vertical eddy diffusion parameterizations that helps in improving marine stratus prediction off the west coasts of continents. In the following, some details on some of the newer features of the model are provided.

### **3.2.1 *Upgrades to Radiation Parameterization***

The new LW and SW radiation parameterizations are modified GCM versions (v2.3g\_LW and v2.3g\_SW) of the Rapid Radiative Transfer Model (RRTM) (hereafter referred to as RRTMG-LW and RRTMG\_SW) adapted from AER Inc. (e.g. Mlawer et al., 1997; Iacono et al., 2000; Clough et al., 2005). In the CFSR implementation we invoke both

SW and LW radiations at one-hour interval. A maximum-random cloud overlapping scheme is used for cloudy sky radiative transfer, and a climatological aerosol scheme provides the global distribution of aerosol optical depth. Stratospheric volcanic aerosols are also included. For CO<sub>2</sub> amount, historical monthly mean observed data available at WMO Global Atmosphere Watch web site <http://gaw.kishou.go.jp> is used. This data is reprocessed into 15 degree latitude/longitude horizontal grid which is used in the CFSR. In addition to the major radiatively absorbing gases (water vapor, ozone, and CO<sub>2</sub>), many other radiatively active green house gases are also included. They are prescribed as global annual mean values adapted from National Institute of Standards and Technology (NIST).

The RRTMG-LW employs a computationally efficient correlated-k method for radiative transfer calculations. It contains 16 spectral bands with various numbers of quadrature points (g-points) in each of the bands that sum up to a total of 140 g-points. Active gas absorbers include H<sub>2</sub>O, O<sub>3</sub>, CO<sub>2</sub>, CH<sub>4</sub>, N<sub>2</sub>O, O<sub>2</sub>, and four types of halocarbons (CFCs). The RRTMG-SW contains 14 spectral bands with various numbers of g-points in each of the bands to a total of 112. It uses a fast two-stream radiative transfer scheme, and includes many absorbing gases (H<sub>2</sub>O, O<sub>3</sub>, CO<sub>2</sub>, CH<sub>4</sub>, N<sub>2</sub>O, O<sub>2</sub>).

### ***3.2.2 Enthalpy as prognostic variable***

The implementation of specific enthalpy,  $C_p T$ , (where  $C_p$  is the specific heat at constant pressure and  $T$  is the temperature) as the thermodynamic prognostic variable, instead of the traditional virtual temperature, is given in Juang (2009). With this approach, both the gas constant  $R$  and specific heat at constant pressure  $C_p$  of air are defined as weighed mean of  $R$  and  $C_p$  for individual gases in the air. Thus this formulation takes into account the variation of  $R$  and  $C_p$  in space and time. However, current physical parameterizations do not take advantage of variation of  $C_p$  and  $R$ . The thermodynamic equation is more accurate when formulated in terms of enthalpy and this is very important for atmospheric models that reach up to 600 km in altitude, where density is very small.

### ***3.2.3 Modification to shallow convection and vertical diffusion***

Both current operational CFS and GFS have very poor prediction of marine stratus near the west coast of major continents. This poor performance may be attributed to the parameterization of shallow convection and background vertical eddy diffusion (Moorthi et al, 2009). The shallow convection parameterization follows Tiedtke et al. (1983). It is applied wherever deep convection parameterization is not active. In this scheme, the highest positively buoyant level below the  $0.7 * P_s$  (where  $P_s$  is surface pressure) level for a test parcel from the top of the surface layer is defined as the shallow convection cloud top. The cloud base is the lifting condensation level (LCL) for the same test parcel. Enhanced vertical eddy diffusion is applied to temperature and specific humidity within this cloud layer. The diffusion coefficients are prescribed with a maximum value of  $5 \text{ m}^2\text{s}^{-1}$  near the center and approaching zero near the edges. The GFS also includes a background eddy vertical diffusion to enhance mixing close to the surface, where eddy diffusion calculated by the PBL parameterization is considered inadequate, particularly in the stable regime. The diffusion coefficient decreases exponentially with pressure, with the surface value set to  $1.0 \text{ m}^2\text{s}^{-1}$ . The combination of shallow convection and background diffusion reduce the strength of low level inversions off the west coast of continents thus reducing the prediction of marine stratus. In the version of the model used in CFSR, two simple modifications (Moorthi et. al., 2009) are used to improve the prediction of marine stratus. The modifications are 1) limit the shallow convection top to be below the low-level inversion when condition for cloud top entrainment instability (Deardorf, 1980; Randall 1980) is not satisfied and 2) set background vertical diffusion to zero above low-level inversions. Combination of these two modifications, although not perfect, does improve the prediction of marine stratus.

#### **4. *The Ocean***

##### **4.1 *The Sea Surface Temperature Analysis***

Two daily sea surface temperature (SST) analysis products have been developed using optimum interpolation (OI). Both products have a spatial grid resolution of  $1/4^\circ$ . One product uses Advanced Very High Resolution Radiometer (AVHRR) infrared satellite

SST data. The other uses AVHRR and Advanced Microwave Scanning Radiometer (AMSR) on the NASA Earth Observing System satellite SST data. Both products also use in situ data from ships and buoys and include a large-scale adjustment of satellite biases with respect to the in situ data. The in situ and corrected satellite data are analyzed using an OI procedure. The correlation scales range from 50-200 km with smaller scales in higher latitudes (especially in western boundary current regions) and larger scales in the tropics. Because of AMSR's near all-weather coverage, there is an increase in OI signal variance when AMSR is added to AVHRR.

The AVHRR-only product uses Pathfinder AVHRR data (currently available from September 1981 through December 2005) and operational AVHRR data for 2006 onwards. The AMSR+AVHRR product begins with the start of AMSR data in June 2002. In this product, the primary AVHRR contribution is in regions near land where AMSR is not available. However, in cloud-free regions, use of both infrared and microwave instruments can reduce systematic biases because their error characteristics are independent.

In the CFS reanalysis the more accurate combined SST product was used when available. Thus, the AVHRR-only product was used from November 1981 through May 2002 and replaced by the AMSR+AVHRR product from June 2002 onward. Both products are available as version 1 and 2. Version 1 of the daily OI is described in Reynolds, et al (2007). This version uses 1 day of satellite and in situ data with satellite bias correction based on 7-days of satellite and in situ data. The changes from version 1 to version 2 (see <http://www.ncdc.noaa.gov/oa/climate/research/sst/oi-daily.php>) are relatively small and primarily consist of additional temporal smoothing. The temporal smoothing includes using 3 consecutive days of satellite and in situ data where the middle day is weighted stronger than the other two days and provides the date of the analysis. The temporal smoothing also includes additional smoothing of the satellite bias corrections using 15 days of satellite and in situ data instead of 7. In addition, ship SSTs are corrected relative to the buoy SSTs by subtracting 0.14°C from all ship observations before they are used to

bias correct the satellite data. Thus, all observations are bias corrected with respect to buoy.

These fields, as prepared by Reynolds at NCDC, do not have temperature values over land. Missing grid points were filled in via interpolation at NCEP to ease conversion of the field to the CFS model grid. Some segments of the CFS reanalysis were run before the full time series of the Version 2 OI daily SST was finalized. For some of these periods, an early release of that product was used.

Ideally, the daily OI products would have been used for the entire CFS reanalysis period. However, these products were not available for the earliest part of the reanalysis: January 1979 through October 1981. For this period, the SST fields prepared for the ERA-40 Project were used. These fields were derived from the UKMO Hadley Center's monthly mean HadISST data set as described in Fiorino (2004).

#### **4.2    *The Ocean Model (MOM4)***

The oceanic component is the GFDL Modular Ocean Model version 4p0d (MOM4p0d) (Griffies et al. 2004), which is a finite difference version of the ocean primitive equations configured under the Boussinesq and hydrostatic approximations. The model uses the tripolar grid developed by Murray (1996). Northward of 65°N it uses a rotated bipolar grid that places two poles over land eliminating the singularity at the North Pole, while southward of 65°N it uses conventional spherical coordinates. The horizontal layout is a staggered Arakawa B grid and the  $z$  coordinate is used in the vertical. The ocean surface boundary is computed as an explicit free surface. The zonal resolution is 1/2°. The meridional resolution is 1/4° between 10°S and 10°N, gradually increasing through the tropics becoming 1/2° poleward of 30°S and 30°N. There are 40 layers in the vertical with 27 layers in the upper 400 m, and the bottom depth is approximately 4.5 km. The vertical resolution is 10 m from the surface to the 240-m depth, gradually increasing to about 511 m in the bottom layer. Vertical mixing follows the nonlocal K-profile parameterization of Large et al. (1994). The horizontal mixing of tracers uses the isoneutral method developed

by Gent and McWilliams (1990; see also Griffies et al. 1998). The horizontal mixing of momentum uses the nonlinear scheme of Smagorinsky (see Griffies and Halberg, 2000). We have used a configuration for the MOM4p0d similar to the one we used for the MOM3 in the current operational CFS (Saha et al., 2006), but we have approximately doubled the resolution and the MOM4p0d is fully global with an Arctic Ocean and interactive ice model, where the operational MOM3 is truncated at 64°N and 74°S.

### 4.3 *The Ocean Analysis: GODAS*

The Global Ocean Data Assimilation System (GODAS) at NCEP uses a 3DVAR assimilation scheme that has evolved from a version originally developed by Derber and Rosati (1989). It was first adopted to assimilate temperature data into a Pacific Ocean model and subsequently modified to assimilate other data sets (Behringer et al., 1998; Ji et al., 2000; Behringer and Xue, 2004; Behringer, 2007). In the context of the CFSR the GODAS assimilates temperature and salinity observations by minimizing the cost function:

$$J = \frac{1}{2} (x - x_b)^T B^{-1} (x - x_b) + \frac{1}{2} (y - H(x))^T R^{-1} (y - H(x)),$$

where the first term is the sum of the squared differences between the forecast,  $x_b$ , and the analysis,  $x$ , weighted by the uncertainty of the forecast and the second term is the sum of the squared differences between the observations,  $y$ , and the analysis,  $H(x)$ , weighted by the uncertainty of the observations. The operator,  $H$ , is a linear interpolator between the model grid and the location of the observation. The matrix  $B$  is the background error covariance and the matrix  $R$  is the observational error covariance that includes the representation error, due to the part of the observed field that the model is unable to resolve. The background error covariance matrix,  $B$ , is univariate and thus block diagonal with respect to temperature and salinity. The horizontal covariance is modeled using a diffusion equation approximating a Gaussian function that is stretched in the zonal direction with the stretching being greatest near the equator (Derber and Rosati, 1989; Weaver and Courtier, 2001). The vertical covariance is likewise modeled with a diffusion equation with length scales specified as a function of depth such that at any level the

scale is twice the level thickness. The variance is set proportional to the square-root of the local vertical temperature gradient computed from the forecast. The observational errors are assumed to be uncorrelated, so that  $R$  is a diagonal matrix of the estimated error variances of the observations. The errors assigned to a temperature profile vary with depth according to the square root of the vertical temperature gradient and are scaled to have values between 1°C and 2.5°C. The standard error assigned to a salinity profile is a constant 0.1 psu at all depths.

Temperature and salinity profiles are assimilated at 6-hour intervals using all observations from the previous 10-day interval. The more distant a profile is in time, the less weight it receives in the assimilation. This approach allows relatively sparse ocean observations to have a greater impact on the model state (Derber and Rosati, 1989; Behringer et al. 1998).

An exception to the assimilation procedure is made in the top level of the model (5 meters) where the ocean analysis is accomplished by simple relaxation to gridded fields of temperature and salinity. The purpose is to provide a stronger constraint on the ocean at the interface with the atmosphere. This is done by relaxing the sea surface temperature, every 6 hours, to the daily mean  $\frac{1}{4}^{\circ}$  SST, described in section 4.1.

#### **4.4    *The Ocean Observations***

##### **4.4.1    *Temperature Profiles***

The temperature observations used for assimilation are profiles from expendable bathythermographs (XBTs), from fixed mooring arrays: the Tropical Atmosphere Ocean / Triangle Trans Ocean Buoy Network (TAO/TRITON) (McPhaden et al., 1998) in the Pacific Ocean, the Prediction and Research Moored Array in the Tropical Atlantic (PIRATA) (Bourlès et al, 2008) and the Research Moored Array for African-Asian-Australian Monsoon Analysis and Prediction (RAMA) in the tropical Indian Ocean and from the Argo profiling floats (The Argo Science Team, 2001). The XBT observations collected prior to 1990 have been acquired from the NODC World Ocean Database 1998 (Conkright et al., 1999), while XBTs collected subsequent to 1990 have been acquired

from the Global Temperature-Salinity Profile Project (GTSP) at:

<http://www.nodc.noaa.gov/GTSP>

An important aspect of the collection of temperature profiles that will have impact on any long-term ocean reanalysis is the variability in the distribution of the profiles in time and space over the last 30 years. Figure 32 illustrates, for the profiles used in this analysis, the number of temperature observations available per month as a function of depth between 1980 and the end of 2008. The stratification of colors reflects the penetration depths of different instrument types. The rapid growth of the Argo array after the year 2000 dominates the latter part of the figure.

There also have been significant changes in the geographical distribution of the profiles. Figure 33-34 which show two annual distributions of profiles (1985 and 2008) gives some idea of the scope of these changes. The year 1985 is dominated by XBT observations collected along shipping routes and during research operations; only a few early TAO moorings have appeared in the tropical Pacific. As a result there are large data voids nearly everywhere and the Southern Ocean is largely unobserved. By 2008 the tropical mooring arrays have expanded into all three ocean basins and the global picture is dominated by Argo profiles. Some gaps remain and marginal seas are not sampled by the Argo array, but there is a nearly complete global coverage.

#### **4.4.2 Salinity Profiles**

Prior to the advent of the Argo array subsurface salinity observations were extremely sparse. For that reason we assimilate synthetic salinity profiles in the ocean analysis. Each synthetic profile is constructed from an observed temperature profile and the local climatological T-S correlation based on the World Ocean Database 1998 (Conkright et al., 1999). An important advantage of this approach is that it maintains water mass properties over the course of the reanalysis. The disadvantage is that some salinity variability in the near surface is lost.

#### **4.4.3 Surface Observations**



As indicated above, temperature and salinity in the top level of the model are relaxed to gridded fields of observed sea surface temperature (SST) and sea surface salinity (SSS). The SST data are the daily fields described in section 4.1. The SSS data is a climatological map based on the World Ocean Database 1998 (Conkright et al., 1999). The SST and SSS fields are remapped to the model grid before being used in the analysis.

## **5.      *The Sea Ice***

The accuracy of sea ice coverage is essential for a good reanalysis of the atmosphere and ocean over the polar regions. Global climate modeling studies have shown that sea ice concentrations strongly affect the climate in the Antarctic regions (e.g., Simmonds and Wu (1993)). However, sea ice concentrations were prescribed in the previous NCEP reanalysis (R1 and R2). The CFSR allows sea ice concentration to be predicted in the forecast guess which provides closer sea ice-atmosphere linkages in the data assimilation system. This should improve the description of the main climatic patterns and trends over the polar regions in the CFSR.

### **5.1      *The Sea Ice Concentration Analysis***

The target in this analysis was to produce a global record of surface ice concentration for all points that may freeze anywhere on the globe, on a daily grid of 0.5° latitude-longitude resolution throughout the period of the CFSR. Further, the aim was to produce the best analysis for the given day. This latter goal means that there are discontinuities in the production of the data set, where newer data sets start to be used, and newer method of analysis are used, or where, as for the Great Lakes, a high quality data set ended. Time series analyses are needed to take care of these discontinuities.

From the start of the CFSR through 31 December 1996, the global ice concentrations for most of the globe are simply regridded from [Cavalieri, et al., 1996, updated 2007] (GSFC Ice), except for a) possibly ice-covered regions that lie outside that grid, b) large Canadian lakes, c) the Great Lakes and d) SST-based filtering of erroneous ice in the analysis.

For the Great Lakes, the data used are Assel et al. [2002] from the start of the CFSR

through the end of the data set in spring 2002, and passive microwave thereafter. Those grids are available 1-3 times per week throughout the period when available.

Concentrations were linearly interpolated between the observation dates, and those interpolated values are used here, averaged on to the target 0.5 degree grid from the native 2.55 km Mercator projection. For large lakes in Canada, the Canadian Ice Service [CIS, personal communication] analyses were used for all lakes which were analyzed (initially 34, in November, 1995, increasing to 137 by October, 2007) from November 1995 through October 29, 2007 (start of CFSR processing). Again, concentrations were linearly interpolated in time between observations. From October 30, 2007 to present, the concentrations are the operational NCEP passive microwave sea ice concentration analyses [Grumbine, 2010a].

Large water bodies which may freeze, but which lie outside the domain analyzed in GSFC Ice [Cavalieri et al., 1996, updated 2007], in the period 1978-1996 were analyzed by proxy, as was done for portions of the North American Regional Reanalysis [Mesinger, et al., 2006]. Proxies were generated anew for this work as the domain was much larger, and more data were now available. During the period 1 January 1997 - 30 June 2006 (when both NCEP ice and GSFC ice were available), the NCEP ice analysis was used to identify points (one by one) which lay inside the GSFC ice domain and which had high correlation to concentrations analyzed for points outside the GSFC ice domain -- but still inside the NCEP domain. This includes large lakes like Lake Ladoga and Lake Onega, as well as the Caspian Sea. Due to changes in SST sources for filtering sea ice concentration analyses, the Aral sea, Lakes Balkhash, Hulun Nur, and some others could not be consistently analyzed and were assigned zero ice concentration. This will bias surface heat fluxes high in periods where the real lake is ice covered. Some lakes cannot be observed strictly by modern passive microwave due to land contamination issues, and were assigned land flags in the reanalysis if no other data were available; this includes Lake Athabasca, Lake of the Woods, Lake Nipigon (outside the period of CIS data), Iliamna Lake, and Lake Vanern.

For 1 Jan 1997-29 Feb 2000, the global ice concentration analysis was the NCEP operational ice analysis [Grumbine, 1996] (again, outside the Great Lakes and Canadian Lakes). For 1 March 2000 to the 29 October 2007 the sea ice analysis is the newer NCEP sea ice analysis system [Grumbine, 2010a] applied to archived passive microwave data for DMSP F-13, F-14, and F-15. The old NCEP system was based on the NASA Team1 algorithm [Cavalieri 1992] as was the GSFC ice [Cavalieri et al. 1996, updated 2007]. The newer system is based on the Team2 algorithm [Markus and Cavalieri, 2000]. In the newer NCEP system [Grumbine 2010a], sea ice concentration for each day, for each (northern, southern) hemisphere, is computed by regression of the Stokes-like parameter  $(T85V^2 - T85H^2)^{0.5}$  (where T85V is the 85 GHz brightness temperature at vertical polarization, and T85H is likewise for the horizontal polarization) against the Team2-derived concentration -- for those points that are greater than 100 km from land, and are poleward of 60 degrees latitude. The regression provides an unbiased estimator, and, due to the small footprint of the 85 GHz channel, a higher resolution estimate -- permitting analysis closer to the coast, and inside smaller lakes, than would otherwise be possible with the pure Team2 algorithm. This operational system used the SSMI instruments on DMSP F-13, F-14, and F-15 while those were all available. F-14 stopped providing data in October 2008. F-15 suffered progressively severe corruption of the 22 GHz channel (used for weather filtering, not for sea ice concentration computation) in late 2008 and was removed from NCEP sea ice production 5 March, 2009. AMSRE was added to the operational sea ice system 13 May 2009, using the AMSR-E Team2 algorithm with January, 2009 tie points as described in Markus and Cavalieri [2009]. That date was, unfortunately, concurrent with a data flow outage from AMSRE and data corruption in F-13. This simultaneous failure degraded the quality of the sea ice analysis in May 2009. From June 2009 to present, sufficient data have been available for analysis. February through April was re-analyzed retrospectively using F-13 and AMSRE, so that the CFSR ice is based on those even though AMSRE did not come in to the operational NCEP sea ice analysis until May 13.

The passive microwave weather filters are imperfect, meaning that ice concentrations can be reported from the microwave for reasons other than ice being on the surface, so that a sea surface temperature filter is also used [Grumbine, 1996]. The sea ice concentrations were in general produced before the SST analyses used for the CFSR. Therefore an *a posteriori* filter was used for retrospective analyses through 29 October 2007 [Grumbine, 2009]. The usual SST filtering was also done using AVHRR-Only analysis [Reynolds et al., 2007] for 4 January 1985-10 February 2000. The RTG low resolution analysis [Thiebaux et al., 2003] was used 11 Feb 2001 through 29 October 2007 and RTG high resolution analysis [Gemmill et al., 2007] thereafter. A more detailed document on the CFSR sea ice analysis is in preparation [Grumbine, 2010b].

## **5.2    *The Sea Ice Model***

The sea ice model is from GFDL Sea Ice Simulator with slight modifications. Similar to the ocean model, sea ice model components use a tripolar grid north of 65N, i.e., a grid that has "poles" located in the land masses of northern Canada and northern Russia, in addition to the normal South Pole. There are three layers in the sea ice model, including two equal layers of sea ice and one layer of snow with five categories of sea ice thickness (0-0.1, 0.1-0.3, 0.3-0.7, 0.7-1.1, and a category greater than 1.1 m). Sea ice dynamics is based on Hunke and Dukowicz (1997) using the elastic-viscous-plastic technique to calculate ice internal stress. Ice thermodynamics is based on Winton (2000). It is possible for ice to be transferred between the snow layer and the two ice layers, conservatively, when there is snowfall, evaporation, freezing or melting. When sea ice forms over the ocean it releases latent heat and salt to the ocean. Details can be found in Griffies et al. (2004).

Due to the lack of sea ice observations of thickness and motion covering CFSR period starting 1979, a sea ice merging scheme is used in CFSR to add sea ice concentration into the system. This scheme merges the observed sea ice concentration with the coupled 6-hour forecast guess field to produce a new initial condition at each analysis cycle. During the merging process, quality control is applied to prevent a failure

when there is feedback between the sea ice analysis and the sea surface temperature (SST) analysis. This is done on the model grid after an interpolation is made for SST and sea ice. When SST from the analysis is warmer than 275.3 K, no sea ice is allowed to exist. When the observed sea ice concentration is greater than (or equal to) 15%, sea ice concentration is reset to the observed value in the guess field. When the observed sea ice concentration is less than 15%, sea ice in the guess field is removed. In summer, the melt pond effect on ice albedo is considered in the Arctic. When there are serious problems in the observed sea ice analysis, only the model prediction was used. This occurred during the period May 1-13, 2009.

Since sea ice concentration was “assimilated” in the CFSR, the resulting initial conditions are close to the observations for sea ice concentration and sea ice coverage. Figure 35 shows the sea ice concentrations for September of 1987 and 2007 for the Arctic. Record minimum Arctic sea ice was observed in September 2007. Sea ice extent, defined as the total area with sea ice present (including open water) for which each grid cell has at least 15% sea-ice, is shown in Figure 36 for the Arctic and Antarctic in the CFSR. Large reduction in sea ice is obvious for both summers of 2007 and 2008 over the Arctic. Inter-annual variability is relatively small for both hemispheres as a whole. The overall trend is slightly positive for the Antarctic and negative for the Arctic, which is consistent with previous studies. Due to realistic sea ice distribution and other upgrades, the coupled CFSR has improved many aspects of the analysis of sea ice concentration over the polar regions, compared to the previous R1 and R2 (Wu and Grumbine, 2010, in preparation).

## **6.     *The Coupler***

### **6.1    *Parallel programming model***

The CFS model, which runs on multiple processes with message-passing tools, uses a parallel programming model called MPMD (Multiple Program Multiple Data). There are three programs (executables) in the CFS: the atmospheric model (GFS), the ocean model (MOM4) and the coupler, each having its own data flow. The three programs run independently, but they exchange data as follows: the GFS runs on the atmospheric time

step  $\Delta a$  (3 minutes), MOM4 runs on a *fast* sea ice time step  $\Delta i$  (also 3 mins) for the sea ice model, and a slow ocean time step  $\Delta o$  (30 minutes) for both the ocean model and *slow* sea ice time step, while the coupler runs on a time step  $\Delta c$ , where  $\Delta c = \max(\Delta a, \Delta i)$ . At every coupler time step  $\Delta c$ , the coupler will receive data from both the GFS and MOM4 sea ice model, and send the needed data back to them, respectively. At every ocean time step, besides the data exchanged between GFS and sea ice, the coupler will also receive accumulated variables (fluxes) from the GFS and send them to the ocean model, while receiving data from the ocean and sea ice model and sending them back to the GFS. When these three time steps are the same, the atmosphere is coupled to the sea ice and ocean at every such time step.

## **6.2 Grid architecture**

In the CFS, the atmospheric model, GFS, uses a different grid structure from the grids of the ocean model MOM4, and the MOM4 sea ice model. The CFS makes use of the atmospheric grid in the dummy atmospheric model (ATM) of MOM4. The atmospheric grid in the dummy atmospheric model ATM is set to the same grid as the GFS, but the grid is distributed on different processors; the CFS coupler redistributes the data received from the GFS grid to MOM4 ATM grid. Inside MOM4, the data on the grid of the dummy ATM model is interpolated with a mask to the MOM4 sea ice grid through the exchange grid of the surface boundary layer. When the MOM4 sea ice model needs to exchange data with the ocean model, the data are redistributed between the two models. This structure is shown in Figure 37.

## **6.3 Passing variables**

In the CFS, both instantaneous and accumulated variables are exchanged between the GFS and MOM4 sea ice and ocean models, through the coupler. The instantaneous variables from the GFS to the sea ice model include downward short wave and long wave radiation, bottom model layer temperature, wind, humidity, pressure, and snowfall. The accumulated variables from the GFS to the ocean are net downward short and long wave radiation, sensible and latent heat flux, wind stress and precipitation. The variables sent

from the sea ice/ocean back to the atmosphere include sea surface temperature, sea ice fraction and thickness and snow depth.

## **7.     *The Land***

### **7.1     *The Precipitation Analysis***

Two sets of global precipitation analyses are used in the CFSR land surface analysis. The pentad data set of CPC Merged Analysis of Precipitation (CMAP, Xie and Arkin 1997) defines 5-day mean precipitation at a 2.5° latitude / longitude grid over the globe by merging information derived from gauge observations, as well as satellite observations in infrared and passive microwave channels. The other data set used is the CPC unified global daily gauge analysis, constructed on a 0.5° latitude / longitude over the global land through the interpolation of quality controlled rain gauge reports from ~30,000 stations collected from the Global Telecommunication System (GTS) and many other national and international collections (Xie et al. 2010). The optimal interpolation (OI) algorithm of Xie et al. (2007) is employed to partially account for the orographic enhancements in precipitation. In addition to the analyzed values of precipitation, the number of reporting stations and the ending time of daily accumulation are also included in the data set. Both analyses are generated for the entire CFSR analysis period from 1979 to the present, and continue to be updated on a real-time basis.

### **7.2     *The Snow Analysis***

Snow liquid equivalent depth was updated using analysis data from the Air Force Weather Agency's SNODEP model (Kopp et al. 1996) and the NESDIS Interactive Multisensor Snow and Ice Mapping System (IMS) (Helfrich et al. 2007). SNODEP uses in situ observations, an SSM/I-based detection algorithm, and its own climatology to produce a global analysis of physical snow depth, once per day at 47 km resolution. Analysts may further adjust the analysis. SNODEP has been operational since 1975 and the data is available for the entire reanalysis period. The IMS data is a manually generated northern hemisphere snow cover analysis produced once per day. Analysts use surface data, geostationary and polar orbiting imagery, and microwave-based detection

algorithms to determine whether an area is either snow covered or snow free. IMS data is available at 23 km resolution starting February 1997, and at 4 km resolution starting February 2004.

### **7.3     *The Land Surface Analysis***

The land surface model (LSM) used in CFSR is the Noah LSM (Ek et al., 2003), which was implemented in the NCEP Global Forecast System (GFS) for operational medium-range weather forecast in 2005. Within CFSR, Noah is implemented in both the fully-coupled land-atmosphere-ocean model to make the first guess land-atmosphere simulation, and in the semi-coupled CFSR Global Land Data Assimilation System (GLDAS) to perform land surface analysis. The semi-coupled GLDAS is forced with the CFSR atmospheric data assimilation output and observed precipitation. GLDAS interacts with the reanalysis once per day, instead of every time step as in its fully-coupled counterpart. The NASA Land Information System infrastructure (LIS, Peters-Lidard et al., 2007) is employed to execute CFSR/GLDAS. This semi-coupled GLDAS/LIS has been configured with the identical setup as in the fully-coupled CFS/Noah, including the same T382 global gaussian grid specification, land-sea mask, terrain height, soil and vegetation classes, and soil and vegetation parameters.

Compared to R1 and R2, this CFSR GLDAS/LIS uses observed global precipitation analyses as direct forcing to the land surface analysis, rather than the typical reanalysis approach of using precipitation from the assimilating background atmospheric model (R1), or using observed precipitation to “nudge” soil moisture (R2). The CPC pentad CMAP analysis and daily gauge analysis are used. Considering global gauge distribution and the strength of the satellite-based precipitation analysis, an optimal precipitation forcing is generated by blending the two precipitation analyses with the CFSR background 6-hourly GDAS precipitation. The blending function is latitude-dependent that favors the satellite-based CMAP analysis in the tropics, the gauge analysis in mid-latitudes, and the model precipitation in high-latitudes. Moreover, an even heavier



weight is assigned to the gauge analysis in regions of dense gauge network, namely, North America, Western Europe, and Australia.

Every LSM is characterized by a specific equilibrium land surface climatology. That equilibrium can be quite different from model to model. The spin-up time required to drive an LSM to its equilibrium is much longer than for the troposphere. Experiments have estimated that at least 3 to 5 years might be required to spin-up the 4-layer CFSR Noah land surface states, if initialized with a previous global reanalysis in which a different LSM (for instance, the 2-layer Oregon State University OSU/LSM (Pan and Mahrt, 1987) in R1/R2) was used. Since the same Noah LSM has been included in the operational GFS since 2005, the 2-year (2006 and 2007) averaged GFS land surface states for each given calendar day of the start dates of the CFSR streams was used as land initial conditions. An additional 12-month spin-up period was executed prior to the CFSR production.

The CFSR GLDAS/LIS is executed once over each 24-hour land surface analysis cycle at 0Z model time, instead of 6-hour cycles of the atmospheric analysis GDAS and the oceanic analysis GODAS. The reason is that GLDAS/LIS is anchored to the *daily* gauge precipitation analysis; hence the gain of executing on a 6-hour cycle is limited. The Noah simulation is made for the past 24 hours using GDAS atmospheric forcing and the blended precipitation forcing. After completion of the 24-hour execution, the simulated soil moisture and soil temperature of all four Noah soil layers are inserted into the CFSR 0Z restart file (the so called “surface file”) as the land surface initial conditions for the next CFSR 0Z cycle.

The IMS and SNODEP data were used to produce daily analyses of physical snow depth on the model physics grid over land. The data were horizontally interpolated using a “budget” method (Accadia et al. 2003) in order to preserve the total water volume. In the southern hemisphere, and globally prior to February 1997, these analyses were created solely from the SNODEP data. In the northern hemisphere starting February 1997, a combination of SNODEP and IMS was used. IMS data was introduced because it more

accurately depicts snow cover compared to SNODEP, especially along mountain ridges (because of its higher resolution). Therefore, in regions where the IMS and SNODEP analyses did not agree, the IMS determined whether there was snow or not in the daily analysis. More specifically, if the IMS indicated snow cover, the analyzed depth was set to 2.5 cm or the SNODEP value, whichever was greater. And if IMS indicated a region was snow free, the analyzed depth was set to zero.

The model snow was updated at 0Z by comparing the first guess to the daily analysis. The analyzed physical depth was converted to liquid equivalent depth using a 10:1 ratio. If the first guess depth was greater than twice (or less than half) the analyzed depth, then the model depth was set to twice (half) the analyzed value. Otherwise, the model snow was not modified. In contrast to directly replacing the model snow with the analysis, this method results in a smoother evolution of the snow pack and reduces the artificial addition of water when the land surface model erroneously melts the snow too quickly. Daily analyses were not available for 65 days due to missing IMS or SNODEP data. On these days, the model snow was simply cycled. GLDAS/LIS also updates its snow fields (snow liquid equivalent and physical depth) to the values of the snow analysis at 00Z. Figure 38-39 shows the CFSR resulting global 2-meter volumetric soil moisture climatology for May and November (1980-2008) respectively. It is consistent with our current knowledge about large-scale soil moisture climatology of wet and dry regions. The contrast between May and November also illustrates the seasonal variation corresponding to precipitating and drying seasons of various regions, for instance, North America, Amazon, and India.

## **8.      *The Execution***

To ensure the project would be completed in a 2-year period, the CFSR was produced by running 6 simultaneous streams of analyses, covering the following periods:

- Stream 1:      1 Dec 1978 to 31 Dec 1986
- Stream 2:      1 Nov 1985 to 31 Dec 1989
- Stream 5:      1 Jan 1989 to 31 Dec 1994

Stream 6: 1 Jan 1994 to 31 Mar 1999

Stream 3: 1 Apr 1998 to 31 Mar 2005

Stream 4: 1 Apr 2004 to 31 Dec 2009

As can be seen, there is a full 1-year overlap between the streams to address spin up issues concerning the deep ocean, the upper stratosphere and the deep soil. Thus the entire CFSR covers 31 years (1979-2009) + 5 overlap years

Figure 40 shows the CFSR execution of one day of reanalysis, which can be itemized as follows:

- Atmospheric T382L64 (GSI) Analysis is made at 0Z, 6Z, 12Z and 18Z, using coupled 9-hour guess forecast;
- Ocean and Sea Ice Analysis (GODAS with MOM4) is made at 0Z, 6Z, 12Z and 18Z, using same 9-hour coupled guess forecast;
- From each of the 4 cycles, a 9-hour coupled guess forecast (GFS at T382L64) is made with 30-minute coupling to the ocean (MOM4);
- Land (GLDAS) Analysis using observed precipitation with Noah Land model is made only at 0Z;
- Coupled 5-day forecast, from every 0Z initial condition, is made with an identical, but reduced horizontal resolution (T126L64) version of the atmosphere, for a sanity check.

Before the actual production phase of the CFSR, a “lite” version (CFSR-Lite) was carried out to sweep through the entire data inventory. This was done with the uncoupled atmospheric model of the CFSR at T62L64 resolution. Each year was a single stream.

## **9. *Preliminary Results of the Atmospheric Analysis***

### **9.1 *Medium Range Forecast Skill***

An integral part of the CFSR job suite was a once daily at 0000 GMT 120 hour medium range global prediction run at the future CFS resolution of T126L64. The primary

purpose of the forecast was to monitor the forecast scores as a measure of quality of the CFSR analyses. The standard measure of medium range forecast skill is the hemispheric 500 hPa anomaly correlation (AC), taken for both hemispheres (NH, SH). The time series of scores could be assessed against those compiled by the operational GFS, R1/CDAS, and the set of R2/CFS hindcasts and the CFSR-Lite. Figure 41 is an example of the yearly NH time series comparing the operational GFS, the CFSR, and the CFSR-Lite for 2006. Note that the yearly average of CFSR, 0.826, is only slightly smaller than that for the operational GFS, which had the advantage of running at the full resolution of T382L64.

#### **9.1.1 *Global Reanalysis Forecast Skill Comparisons***

The daily 120 hour forecast scores for each hemisphere have been averaged for each year in the period 1979-2008 for the CFSR, CFSR-Lite (through 2006), the operational GFS, CFS R2 hindcasts (1981-2006) and CDAS/R1. The NH scores are on the right of Figure 42 and the SH scores are on the left. The configuration of the systems was as follows:

CFSR: T382L64 (GSI)

CFSR-Lite: T62L64 (GSI)

CFS R2 Hindcasts: CFS: T62L64, R2: T62L28 (SSI)

CDAS/R1: T62L28 (SSI)

GFS: 1984: R40L12; 1985: R40L18; 1987: T80L18; 1991: T126L28; 2000: T170L42; 2002: T254L64; 2005: T382L64.

Please note the resolution changes of the GFS may include changes to the forecast model and/or data assimilation. A history of changes to the NCEP GFS/GDAS may be found at: [http://www.emc.ncep.noaa.gov/gmb/STATS/html/model\\_changes.html](http://www.emc.ncep.noaa.gov/gmb/STATS/html/model_changes.html)

#### **9.1.2 *Northern Hemisphere (NH) Results***

The CDAS/R1 scores are remarkable steady over the past 30 years, with very slight improvement since 2000, which may be related to the automation of rawinsondes, the debut of ACARS ascent and descent profiles, and the improvement of ATOVS retrievals over those of TOVS. Note the comparable scores for CDAS/R1 and the operational GFS

in the mid 1990's, the period when R1 was assembled and run based on the GFS of that era. The CFS R2 hindcast scores are a slight improvement over CDAS/R1, which may be attributed to the increase of vertical resolution from 28 to 64 layers in the CFS predictions and ex post facto hindcast bias correction. CFSR-Lite has a modest, but clear, improvement over the CFS R2 hindcasts, a result of multiple factors: 1) a vertical resolution increase of the assimilation system, from 28 to 64 layers, 2) direct assimilation of radiances replacing retrievals, 3) GSI vs. SSI, and 4) forecast model improvements between 2003 and 2007. The CFSR represents a quantum leap improvement over the lower resolution systems for the entire period 1979-2008, and demonstrates that the level of skill of the comparable operational GFS has been maintained since 2000. The dramatic jump in CFSR skill over the CFSR-Lite emphasizes the importance of horizontal resolution in global data assimilation. The slight improvement of the operational GFS over the CFSR, since the GFS resolution increased to T382L64 in 2005, may well reflect the importance of resolution in the medium range prediction model. The upward slope of the CFSR scores reflects both the conventional data improvements noted for above, and the TOVS to ATOVS radiance transition.

### **9.1.3 *Southern Hemisphere (SH) Results***

The most noticeable difference between SH and NH scores (Fig 42) is the slope of the increase in skill over the period 1979-2008 compared to those for the NH. The SH scores may be viewed partly as a proxy for oceanic observing systems (e.g. satellite winds and radiances), and the NH a proxy for land based systems (e.g. rawinsondes and ACARS profiles). The CDAS/R1 and GFS scores did not consistently sustain above the 0.60 skillful level until the early 1990's. As in the NH, CFS R2 hindcasts only marginally improved over CDAS/R1. The quantum leap in skill came with CFSR-Lite scores, which put an exclamation point on the importance of the direct assimilation of radiances, with a comparable level of skill maintained throughout the TOVS period of 1979-1998. The superiority of the ATOVS instruments over the TOVS is noted by the jump in CFSR-Lite scores since 2000. The CFSR scores are modestly better than the

CFSR-Lite for the TOVS period, but dramatically better in the ATOVS period, a suggestion that the higher resolution CFSR and GFS assimilation systems are exploiting information in the ATOVS systems not available from TOVS. Note that the NH and SH CFSR and GFS scores at the end of the period are nearly comparable, an indication that the medium range skill of the prediction of synoptic scale features is now comparable over oceans and continents.

## **9.2    *Atmospheric Mass***

Monitoring atmospheric mass as part of a reanalysis effort is now common, see Fig. 10 in Uppala et al (2005) and Fig. 1 in Trenberth and Smith(2005), which compared traces of surface pressure, precipitable water and dry pressure in ERA40, ERA15 and R1. Compared to previous reanalyses, the new CFSR is very well behaved in terms of the atmospheric mass or, very nearly equivalent, the surface pressure. One sanity check is to study the time variation of the ‘dry pressure’, which should almost always be conserved. The dry component of surface pressure is defined locally by subtracting the precipitable water (in all its forms) from the total surface pressure, converted to pressure units. Indeed, the global mean of dry pressure is almost constant, with a standard deviation of 0.10 hPa around its mean value of 983.01 hPa over the 1979-present period (not shown). The global mean of the total pressure varies somewhat more, a standard deviation of 0.16 hPa around its mean of 985.39 hPa, since the water content of the atmosphere is free to vary. However, even this variation is less than 1 hPa from maximum to minimum in the global mean. In earlier reanalyses (Trenberth and Smith 2005), these variations were much larger and a consistency between total pressure and precipitable water was lacking.

The mass balance of the atmospheric water component, with its input by evaporation (E) and output by precipitation (P), remains somewhat worrisome even in the CFSR. Global average monthly mean precipitation (P), evaporation (E) and E-P are shown in Figure 43. The global mean P is always larger than E (by a non-negligible few tenths of a mm/day) and this imbalance grows around 1998, probably related to the ingest of new data systems, like AMSU. The decrease of global mean E-P after 1998 appears to

be due to a change over the oceans (Figure. 43 b and c) and results from both an increase in precipitation (Figure. 43a) and a decrease in evaporation (Figure. 43b). Another feature is an increase in the amplitude of the seasonal cycle of E-P over land, after 1999 (Figure. 43c, blue curve), which appears to be due to a weaker seasonal amplitude in precipitation (Figure. 43a). Further analyses are required to investigate the causes of these features in the CFSR. Assimilation of observations could, in principle, violate the governing equations, and the hydrological imbalance in the CFSR is one prominent example of this violation.

### **9.3    *Atmospheric Tides***

We now discuss the 2D atmospheric mass distribution in some detail here, namely the atmospheric thermal (solar radiation induced) tides in the CFSR. The previous reanalyses have had many types of users. Among them, geodesists and oceanographers (Ponte and Ray 2002; Ray and Ponte 2003; Velicogna et al 2001; de Viron et al 1999), who require global atmospheric surface pressure, especially atmospheric tides, for their research and real-time applications. Compared to R1, the new CFSR offers several very significant advances for such users. First, and perhaps foremost, is the availability of hourly output. Figure 44 gives an example of the solar tides as analyzed by CFSR for March 1998 (an arbitrarily picked month). The 24 global maps in one figure give a compact display of the tides. Each map is obtained from data available on the model T382 Gaussian grid. In the red (blue) areas, the pressure is higher (lower) than the daily mean. Units are hPa. Time starts in the upper left (00Z) corner, then proceeds down to 3Z, continues at the top of the next column to 4Z and so on, to the end of the diurnal cycle at 23Z. The data at 0, 6, 12 and 18 Z are from the actual analysis made at these times, while the data in between these cycles is from the coupled guess forecast. One can clearly see the red and blue areas propagate westward around the earth in one day. A cross-section along the equator shows a dominant wave number two pattern, corresponding to the semi-diurnal solar oscillation. This type of map is available for each month from 1979-present, at the CFSR website at: <http://cfs.ncep.noaa.gov/cfsreanl>

A simple operation, making a monthly mean for each hour of the day, is enough to filter out most weather and can produce climatological features, such as the tides. Applications to both geodesy and oceanography require knowledge about atmospheric tides at high temporal frequency, such as hourly intervals or better. However, most reanalyses to-date have output only every six hours – imagine Figure 44 with only 0Z, 6Z, 12Z and 18Z entries present. Previously, interpolation methods had to be devised (Van den Dool et al, 1997) to extract hourly tides from six-hourly R1 data. While these methods worked semi-satisfactorily, it is better to obtain hourly data directly from the assimilating model. Thanks to larger data storage devices, this can be done now. Nevertheless, users should keep in mind that only the 0, 6, 12 and 18Z fields are actual analyses, while the in-between hourly data are model forecasts. The second major advance is that the tides appear to be far more accurate in terms of their amplitude (the phase error was never a big problem). R1 had a semidiurnal variation with an amplitude about 40% too strong (see Table 1 in Dai and Wang (1999)), compared to the gold-standard, the global station analysis by Haurwitz and Cowley(1973). This amplitude error has disappeared completely in the CFSR. A third advance is a far higher spatial resolution of the CFSR.

#### **9.4 Accuracy of the Surface Pressure Analysis**

The accuracy of the surface pressure analysis in the CFSR, as expressed by computing error bars on the analyzed surface pressure fields, is difficult to judge directly. It is too expensive to conduct logical cross-validation experiments, such as making a reanalysis with say, 10% of the data being withheld in turn, to judge accuracy. Instead, we discuss the root-mean-square difference of the 6 hour forecast guess against the surface observations (wherever they are), also called background error. Since error growth is limited in 6 hours, this measure is only slightly higher than the accuracy of the analysis. Figure 45 shows annually compiled fit-to-obs statistics for SH ocean (blue) and NH land (red) for 1979-2008. The accuracy of the surface pressure analysis over the SH ocean has steadily improved, from an error of 2 hPa in 1979 to only about 1.0 hPa at present. The improvement is gradual, and relates to a gradual increase in observations (and not just



surface pressure observations) over the SH oceans, which were sparsely observed before there were satellites. At the end of the period, the error in the analysis over the SH oceans is on par with that over the NH land. Figure 45 shows that NH land has an error of about 1.1 hPa from 1979 to 1996, without the gradual improvement seen in the SH. Then, suddenly in May/June 1997, a sudden large increase in METAR data (see Figure 7) helped reduce the error in the NH to below 1 hPa. In general, one may conclude that the accuracy of surface pressure analyses in the hemispheres (whether it be land or ocean), is getting closer. This has been noticed when real time operational 5 day anomaly correlation scores for the two hemispheres became comparable, after a long history of lagging scores in the SH (see red lines for the GFS in Figure 42).

The better performance discussed above, relative to R1, has many reasons. Among them are improved observational datasets, new data sets (not previously used), improved ongoing monitoring during the execution of the CFSR, and about 15 years of advances in the forecast models and data assimilation methods. The higher resolution of the CFSR also helped in allowing more data to be assimilated over complex terrain. A more thorough analysis of atmospheric mass in the CFSR is forthcoming in a separate paper (van den Dool and Saha, 2010).

### **9.5    *The Upper Atmosphere***

The CFSR is an improvement upon R1 and R2 due to increased horizontal (T382 vs. T62) and vertical resolution (L64 vs. L28). The L28 model has only 8 layers above 100 hPa, with none being above 1 hPa. By contrast, the L64 model has 22 layers above 100 hPa, with 2 layers being above 1 hPa. Because of the vertical limitations, R1 and R2 only produce usable analyses up to 10 hPa. However, the CFSR can produce useful analyses up to 1 hPa.

Another notable difference between R1/R2 and the CFSR is that the CFSR GSI uses satellite radiances rather than derived temperature or moisture profiles. This allows the GSI greater freedom to generate adjustments to the temperature, moisture, and ozone fields to best match the observed radiances. The use of satellite radiances from ozone

sensitive channels mandates the need for quality ozone profile information. The CFSR assimilates the version 8 SBUV profile and total ozone retrievals (Flynn et al., 2009) from NIMBUS-7 and SBUV/2 profile and total ozone retrievals from NOAA 9, 11, 14, 16, 17, and 18. The ozone mixing ratios and total ozone values have not been adjusted to account for biases from one satellite to the next.

Previously, it was shown how the number of radiosonde reports per year and the number of sondes reporting above 100 hPa, increased in the late 1990's. The GSI had difficulty determining equatorial winds due to the low density and quality of radiosonde observations prior to 1998. Hence, the CFSR had difficulty analyzing a proper QBO and Semi Annual Oscillation (SAO) wind signal before that time. ERA-40 winds above 20 hPa from 20S to 20N were assimilated until 1998, when AMSU data became available, and radiosonde coverage was better. Figure 46 shows the global mean temperature anomalies from 1000 to 1 hPa from January 1979 through May 2009. Issues of using the SSU data from 1979 to 1998 have been evident in both the ERA-40 and JRA-25 temperature anomalies (Onogi, et al., 2007). Extensive collaboration between the JCSDA, NASA and ECMWF resulted in minimizing the effects of the SSU CO<sub>2</sub> pressure changes, and the frequent loss of satellite channels. However, in the CFSR assimilation process, all 3 of the SSU channels were subject to bias correction. Unfortunately, these corrections, of the highest channel (3) in particular, resulted in feedback which produced a gradual warming of almost 10°C in the upper atmosphere over time. The end and beginning of each stream is quite evident as warm anomalies are transitioned to cold biases. The upper atmosphere anomalies are greatest at the lowest pressures and reach down to 30 hPa. These anomalies decrease in depth after 1998, as the CFSR progresses to the present. Positive temperature anomalies from the El Chicon (April, 1982) and the Mt. Pinatubo (June, 1991) volcanic eruptions are apparent between 100 and 30 hPa, and last about two years. Evident in the troposphere is a gradual change from negative anomalies to positive anomalies between 700 and 200 hPa.

Figure 47 shows the equatorial (5°S to 5°N) zonal mean U wind component time series from 1979 to 2009. Evident are the annual equatorial easterlies in the troposphere. The QBO variability of the u component is evident in the lower stratosphere while the SAO is evident in the upper stratosphere to lower mesosphere.

Figure 48 shows the difference between the zonal monthly mean total ozone values in the CFSR and that of the Climate Prediction Center's Cohesive SBUV(/2) Total Ozone Dataset (CSTOD). Most of the differences are in the +/- 4 DU range, which is slightly larger than 1 percent. As shown by the aid of the dashed lines, some differences start and stop as different satellites are used for the CSTOD. These differences can be explained by the fact that the CSTOD makes latitudinal bias adjustments from one satellite to the next, whereas there are no adjustments to the SBUV(/2) data used in the CFSR. It has also been discovered that the ozone mixing ratio background errors in the CFSR were set too large in the upper stratosphere. The result is that SBUV(/2) ozone mixing ratio observations did not alter the CFSR's first guess above 10 hPa. Although the integrated effect upon the total ozone is small, this does account for the differences in the southern high latitudes between 1979 and 1994.

Future reanalyses would benefit from not bias adjusting the SSU radiances and having fewer streams. Improvements to the GSI to handle sparse data in the tropics would result in a better analysis of the QBO and SAO winds. Ozone is one of the few parameters that the GSI assimilates as a physical quantity rather than as observed radiances. Inter-satellite bias adjustments must come from improved versions of the SBUV(/2) ozone data record. How the GSI uses the ozone data and alters its first guess must be looked into further.

## **10. *Preliminary Results from the Ocean***

### **10.1 *Precipitation-SST relationship***

Previous studies have shown that there is a close relationship between precipitation and SST on intraseasonal time scales in the tropical Indian Ocean and in the Western Pacific Ocean.

In Figure 49, we show the temporal lag-correlation coefficient between precipitation and SST in the tropical Western Pacific region in two generations of NCEP reanalysis data. Data for the boreal winter (Nov-Apr) over the period 1979-2008 are band-pass filtered for 20-100 days after removing the climatological mean.

It shows clearly that the precipitation-SST relationship in CFSR is consistent with observational data: At lag 0, precipitation has a weak negative correlation with SST. The positive correlation of SST and precipitation gradually increases with a warming of the SST (negative lags on the horizontal axis in the figure), reaching a maximum at lag day - 7. On the other hand, the cooling of SST by precipitation reaches a maximum at lag day 5 (positive lags on the horizontal axis in the figure). In R1 and R2, the increase in precipitation due to warming of the SST is too quick, which implies a lack of coupling in the data-assimilation system or a problem with using an observed SST which was a weekly-averaged product. In the coupled CFSR, this deficiency is largely corrected. These results are consistent with the observational study of Woolnough et al. (2000) and the coupled vs uncoupled model comparison studies of Pegion and Kirtman (2008); and Fu and Wang (2004).

## **10.2 *The Oceanic Component***

In this subsection, we present a few results from the oceanic component of the CFSR, and compare them to observations.

Equatorial cross-sections of temperature are shown in Figure 50 for the CFSR (top) for the years 1982-2008 and for observations from the World Ocean Atlas (WOA) (bottom) obtained from Conkright et al. (1999). The 20°C isotherm, which, at the equator, is close to the maximum vertical temperature gradient (plotted as a black line) slopes upward toward the east, from a depth of approximately 150 m in the western equatorial Pacific Ocean to approximately 20m depth in the equatorial cold tongue region. While the depths of the 20°C isotherms in the two sections are in good agreement, the thermocline in the

CFSR is more diffuse, such that that CFSR section is cooler than the WOA section above the 20°C and warmer below.

The corresponding equatorial cross-section of the zonal velocity in the CFSR is shown in Figure 51. The isopleths slope upward toward the east in the equatorial Atlantic, Pacific and Indian Oceans. The core of the undercurrent is approximately centered on the 20°C isotherm and the maximum mean velocity is about 0.85m/s at 130°W in the equatorial Pacific.

The equatorial temperatures and zonal velocities are also compared with the TAO mooring data in Figure 52. Vertical profiles of the subsurface temperature (top panels) and zonal velocities (bottom panels) are shown for four locations in the equatorial Pacific Ocean for the CFSR (red line) and for the TAO observations (black line). The CFSR temperature profiles are very close to the TAO observations. This is an expected result since TAO observations are assimilated by the CFSR. One noticeable difference is that the CFSR temperatures are approximately 1°C colder than the TAO temperatures below the thermocline in the western equatorial Pacific. The core depth of the undercurrent in the CFSR is in good agreement with the TAO observations in the eastern part of the basin, although the magnitude of the undercurrent in the CFSR is about 10-15% less than observed. The CFSR also has difficulty capturing the mean westward flow in the central basin that is seen in the TAO data.

The vertically averaged temperature is plotted in Figure 53 for CFSR for the years: 1979-2008, and for observations from the World Ocean Atlas (Conkright et al., 1999). Note that the plots show remarkable agreement, with the main difference being that the CFSR is approximately 0.5°C colder than the observations between 30S-30N.

In Figure 54, zonal and meridional velocities from the first level (5m) of the CFSR are compared with the same fields derived from surface drifters drogued at 15m (Lumpkin and Pazos, 2006). In making this comparison it should be kept in mind that the CFSR maps were made from averages of Eulerian velocities on the model grid, while the drifter maps were constructed from pseudo-Lagrangian motion of drifters non-uniformly

distributed in time and space. Despite this difference the patterns in the two sets of maps are remarkably similar. The drifter velocities are distinctly stronger in the Antarctic Circumpolar Current, the western boundary currents and the equatorial zone of the Indian Ocean. The zonal velocities in the equatorial zones of the Atlantic and Pacific are similar in the CFSR and drifter maps although the drifter velocities are more divergent in the cold tongue region of the eastern Pacific.

In Figure 55, the first two EOFs of the CFSR sea surface height (SSH) are shown at the top and for satellite altimeter data from TOPEX/Jason-1 in the middle. The EOFs were computed for the period 1993-2008, the period for which TOPEX/Jason-1 data are available. The principal components in the bottom panel represent the temporal variability of the two EOFs. The first EOF for CFSR and TOPEX/Jason-1 data represents the warm phase of ENSO, as is indicated by the maximum value in 1997, while the second EOF represents the cold phase of ENSO, as indicated by its maximum value in 1998. The first two EOFs for CFSR SSH are very similar to that for upper ocean heat content (not shown), reflecting the dominance of the thermosteric variability of SSH over other physical processes. While the CFSR and TOPEX/Jason-1 data show remarkable similarity (considering that TOPEX/Jason-1 data are not assimilated in CFSR), the major difference is the double maximum found in the spatial plot of the first EOF when compared to TOPEX/Jason-1. Also, the first EOF represents only 32% of the variability in CFSR compared to 42% for TOPEX/Jason-1. The second EOF is very similar for both CFSR and TOPEX/Jason-1.

## **11. *Concluding Remarks***

In this paper, we have described a new global reanalysis (CFSR) produced at NCEP covering the period 1979 to the present. Sections in this paper describe the observational datasets, forecast models and data assimilation systems used in this endeavor. We also briefly present some preliminary results. The accuracy of this new reanalysis is hard to assess directly, but using the 5-day forecast scores as a measure of the accuracy of initial states, (a reasonable assumption), we conclude that CFSR is considerably more accurate

than the previous global reanalysis made at NCEP in the 1990s. It is more comprehensive, since it includes analyses of both the ocean and sea ice, and it has higher resolution in space and time. The accuracy increases over time, especially in the Southern Hemisphere where the use of satellite radiance data becomes very important.

Many known errors in the observational data ingest and execution of the previous reanalyses was corrected in the CFSR. Many of the input datasets have been improved by years of quality control and by exposure to successive reanalyses at various centers, most notably at ECMWF, NCEP and JMA. Undoubtedly, some errors may still persist and more errors may be discovered in the future, but this is all part of the converging process of making better and better analyses of the Earth's system. A thorough study is also required of the benefits of coupling the atmosphere to the ocean and the sea ice, but at first blush, this aspect appears to have worked very well in the CFSR. The SST/precipitation correlation has improved in the tropics. Some problems related to executing the project in many streams (as is common to speed up the process) still exist. All subcomponents with longer time scales (such as the deep soil, deep ocean and the top of the atmosphere) have discontinuities where one stream ends and the next one begins, this in spite of a full one year overlap between the streams.

Future developments include the following three projects:

1. CFSR was conducted mainly to create initial conditions for the coupled atmosphere-ocean-land-sea ice Reforecasts of the next Climate Forecast System, over the period 1982-present. This project is underway at NCEP. The design of these reforecasts is as follows: From every fifth day in the calendar, there will be four 9-month 'seasonal' forecasts from 0, 6, 12 and 18UTC. From every day, there will also be shorter predictions, one run to the first season (105 days) at 0Z and three runs to 45 days from 6, 12 and 18 UTC. The emphasis on the shorter subseasonal predictions, for the MJO and week 3-6 forecasts, is to bridge the weather-climate gap and was the main reason for a high resolution reanalysis to be conducted. There is consensus that, given a

forecast model at a lower resolution (say T126), the skill of the forecasts benefits from the highest possible resolution of the initial state.

2. Given the pace of model and data assimilation development, we expect a new global reanalysis to be conducted at NCEP, once every 7 years or so. However, there is serious thought being given to immediately conduct a 'light' (with a reduced horizontal resolution of T126) version of the reanalysis just completed. It would be done in a single stream to overcome the discontinuities found in the CFSR for the deep ocean, deep soil and the top of the atmosphere. It is possible that this companion reanalysis will be finished in 1 year, in time for CPC to use it when they change their climate normals to the last 30-year period from 1981-2010.
3. A final activity to be conducted when the Reforecast project is complete is to apply the reanalysis system, as used here, to the historical period 1948-1978. The CFSR is the successor of R2, and when extended back to 1948, will also be the successor of R1. It is possible this will be done in one-stream 'light' mode.

### ***Acknowledgements***

We acknowledge the US Climate Program Office (Reanalysis Grant) for funding part of this project. We thank the staff of EMC, CPC, NCEP Central Operations (NCO) and NESDIS for their various contributions in completing this large endeavor in record time. We acknowledge the Geophysical Fluid Dynamical Laboratory (GFDL) for the ocean and sea ice model used in the CFSR. We thank George Vandenberghe, Carolyn Pasti, Doris Pan, Don Avart, Cameron Shelton and others for their outstanding computer support. We acknowledge the efforts of Augustin Vintzileos (EMC) and Adrian Simmons (ECMWF) to give us the AMMA observations in time for use in the CFSR. The CFSR Science Advisory Board (chaired by Jeff Anderson) gave useful guidance at the start of the project. A very thorough and constructive internal review was provided by Glenn



White (EMC). As with other reanalysis projects at NCEP, this is the collective work of many people from many organizations. We express our gratitude to all of them.

## **Appendix A: *Acronyms***

AER	Atmospheric and Environmental Research
ACARS	Aircraft Communications Addressing and Reporting System
AIRS	Atmospheric InfraRed Sounder
AMI	Active Microwave Instrument
AIRSEV	AIRS EVery spot (9) in an AMSU-a field of view
AMIP	Atmospheric Model Intercomparison Project
AMSR-E	Advanced Scanning Radiometer - EOS
AMSU	Advanced Microwave Sounding Unit
ATOVS	Advanced TIROS Operational Vertical Sounder
ATM	ATmospheric Model
AVHRR	Advanced Very High Resolution Radiometer
BAMS	Bulletin of the American Meteorological Society
BOM	Australian Bureau of Meteorology
BUFR	Binary Universal Form for the Representation of meteorological data
CDAS	Climate Data Assimilation System
CDROM	Compact Disc Read-Only Memory
CFS	Climate Forecast System
CFSR	Climate Forecast System Reanalysis
CFSRR	Climate Forecast System Reanalysis and Reforecasts
CHAMP	CHAllenging Mini-satellite Payload
CIS	Canadian Ice Service
CLASS	Comprehensive Large Array-Data Stewardship System
CMAP	Merged Analysis of Precipitation
COADS	Comprehensive Ocean-Atmosphere Dataset
COSMIC	Constellation Observing System for Meteorology, Ionosphere and Climate
CPC	Climate Prediction Center
CRTM	Community Radiative Transfer Model
DMSP	Defense Meteorological Satellite Program
DOE	Department Of Energy
DVD	Digital Versatile Disc
ECMWF	European Center for Medium Range Weather Forecasts
EMC	Environmental Modeling Center
EnKF	Ensemble Kalman Filter
EOF	Empirical Orthogonal Function
ERA-15	ECMWF 15 Year Reanalysis
ERA-40	ECMWF 40 Year Reanalysis
ERS	European Remote Sensing satellite
ESA	European Space Agency
ESRL	Earth System Research Laboratory
EUMETSAT	European Organization for the Exploitation of Meteorological Satellites
FOTO	First order time interpolation to the observation
GDAS	Global Data Assimilation System
GEOS	Goddard Earth Observing System model
GFDL	Geophysical Fluid Dynamics Laboratory

GFS	Global Forecast System
GMAO	Global Modeling and Assimilation Office
GMS	Geosynchronous Meteorological Satellite
GODAS	Global Ocean Data Assimilation System
GOESND	GOES Sounder
GOESSFV	GOES Sounder Full View
GPSRO	Global Positioning System Radio Occultation
GRIB	Gridded Binary
GSFC	Goddard Space Flight Center
GSI	Gridded Statistical Interpolation
GTS	Global Telecommunication System
GTSP	Global Temperature-Salinity Profile Project
HIRS	High Resolution Infrared Sounder Unit
IASI	Infrared Atmospheric Sounding Interferometer
IMS	Ice Mapping System
JMA	Japanese Meteorological Agency
JRA-25	Japanese 25 Year Global Reanalysis
JCSDA	Joint Center for Satellite Data Assimilation, NCEP/NASA
LDAS	Land Data Assimilation System
Level 1B	Satellite data expressed in sensor counts
Level 1C	Satellite data expressed in radiances (or brightness temps)
MERRA	Modern Era Retrospective-Analysis for Research and Applications
METAR	MÉTéorologique <u>Aviation</u> Régulière (routine aviation weather report)
METEOSAT	Meteorological Satellite
METOP	Meteorological Operational satellite
MHS	Microwave Humidity Sounder
MODIS	Moderate-resolution Imaging Spectroradiometer
MOM	GFDL Modular Ocean Model
MSNET	Mesoscale Network of automated weather stations
MSU	Microwave Sounder Unit
NARR	North American Regional Reanalysis
NASA	National Aeronautical and Space Administration
NCAR	National Center for Atmospheric Research
NCDC	National Climatic Data Center
NCEP	National Centers for Environmental Prediction
NESDIS	National Environmental Satellite, Data and Information Service
NEXRAD	Next generation Radar
NMC	National Meteorological Center
NOAA	National Oceanic and Atmospheric Administration
Noah	Ncep/Oregon state university/Air force/Hydrologic research lab
NODC	National Ocean Data Center
NOMADS	NOAA Operational Model Archive and Distribution System
NRL	Naval Research Laboratory
NSDA	National Space Development Agency of Japan
NWS	National Weather Service (US)

ON	Office Note (ie ON29/124)
PCMDI	Program for Climate Model Diagnosis and Intercomparison
PIRATA	Prediction and Research Moored Array in the Tropical Atlantic
PROFLR	Profiler
QBO	Quasi Biennial Oscillation
QuikSCAT	Quick Scatterometer
R1	Global Reanalysis 1 (NCEP/NCAR Reanalysis)
R2	Global Reanalysis 2 (NCEP-DOE Reanalysis)
RRTM	Rapid Radiative Transfer Model
RTOVS	Revised TIROS Operational Vertical Sounder
RTG	Real Time Global
SAO	Semi Annual Oscillation
SATOB	Satellite Observation (usually means satellite wind)
SNODEP	SNOw DEPth
SSI	Spectral Statistical Interpolation
SSM/I	Special Sensor Microwave Imager
SST	Sea Surface Temperature
SSU	Stratospheric Sounder Unit
SYNOP	Surface SYNOptic report
TAO	Tropical Atmosphere Ocean / Triangle Trans Ocean Buoy Network
TIROS	Television and InfraRed Observation Satellite
TOVS	TIROS Operational Vertical Sounder
UKMO	United Kingdom Meteorological Office
VADWND	Vertical Azimuth Display Wind from NEXRAD
XBT	Expendable BathyThermographs

## Appendix B: *The Data Description*

### Monthly Means

(00Z,06Z,12Z, 18Z and daily averages, analysis with 0-6 hour forecast)

<i>File</i>	<i>Grid</i>	<i>Description</i>	<i>MB/month</i>
<b>FLXF</b>	T382	Surface, radiative fluxes, etc.	3,500
<b>FLXL</b>	T62		105
<b>PGBH</b>	1/2 degree	3-D Pressure level data	11,200
<b>PGBL</b>	2.5 degree		480
<b>OCNH</b>	1/2 degree	3-D Ocean data	2,100
<b>OCNF</b>	1 degree		570
<b>DIABF</b>	1 degree	Diabatic Heating, etc.	2,400
<b>DIABL</b>	2.5 degree		525
<b>IPVH</b>	1/2 degree	3-D Isentropic level data	2,200
<b>IPVL</b>	2.5 degree		100
<b>EGYH</b>	1/2 degree	Energetics, u' v', TKE, etc	7,100
<b>EGYL</b>	2.5 degree		280

### Hourly Timeseries from FLX files (32 parameters)

(analysis with 0-6 hour forecast)

Low resolution T62 Gaussian / High resolution T382 Gaussian

<i>File</i>	<i>Description</i>	<i>(low/high res in MB/month)</i>
<b>DLWSFC</b>	Downward LW at the surface	40 / 605
<b>DSWSFC</b>	Downward SW at the surface	26 / 360
<b>GFLUX</b>	Ground heat flux	7 / 168
<b>ICECON</b>	Ice concentration	5 / 56
<b>ICETHK</b>	Ice thickness	5 / 103
<b>LHTFL</b>	Latent heat flux	12 / 215
<b>PRATE</b>	Precipitation rate	12 / 208
<b>PRESSFC</b>	Surface pressure	18 / 295
<b>PWAT</b>	Precipitable water	14 / 190
<b>Q2M</b>	2m specific humidity	21 / 444
<b>RUNOFF</b>	Ground runoff	9 / 159
<b>SHTFL</b>	Sensible heat flux	12 / 208
<b>SNOHF</b>	Snow phase-change heat flux	3 / 77
<b>SOILM1</b>	Soil Moisture Level 1	8 / 230
<b>SOILM2</b>	Soil Moisture Level 2	8 / 230
<b>SOILM3</b>	Soil Moisture Level 3	8 / 230
<b>SOILM4</b>	Soil Moisture Level 4	8 / 230
<b>SOILT1</b>	Soil Temperature Level 1	11 / 284
<b>SWE</b>	Snow water equivalent	12 / 81
<b>TMAX</b>	Maximum 2m air temperature	11 / 171
<b>TMIN</b>	Minimum 2m air temperature	11 / 171

<b>TMP2M</b>	2m air temperature	24 / 473
<b>TMPHY1</b>	Temperature at hybrid level 1	26 / 470
<b>TMPSFC</b>	Surface temperature	24 / 477
<b>U10M</b>	U at 10m	36 / 633
<b>V10M</b>	V at 10m	(UV combined)
<b>UFLX</b>	U-stress	29 / 491
<b>VFLX</b>	V-stress	(UV combined)
<b>ULWSFC</b>	Upward LW at the surface	10 / 153
<b>ULWTOA</b>	Upward LW at the top	12 / 187
<b>USWSFC</b>	Upward SW at the surface	7 / 101
<b>USWTOA</b>	Upward SW at the top	9 / 154

### **Hourly Timeseries from PGB files (27 parameters)**

(analysis with 0-6 hour forecast)

Low resolution 2.5 degree / High resolution 0.5 degree

<b><i>File</i></b>	<b><i>Description</i></b>	<b><i>(low/high res in MB/month)</i></b>
<b>CHI200</b>	Velocity Potential at 200 hPa	7 / 73
<b>CHI850</b>	Velocity Potential at 200 hPa	7 / 73
<b>PRMSL</b>	Mean Sea Level Pressure	18 / 230
<b>PSI200</b>	Streamfunction at 200 hPa	9 / 82
<b>PSI850</b>	Streamfunction at 850 hPa	9 / 88
<b>Q500</b>	Specific Humidity at 500 hPa	11 / 202
<b>Q700</b>	Specific Humidity at 700 hPa	8 / 150
<b>Q850</b>	Specific Humidity at 850 hPa	8 / 157
<b>Q925</b>	Specific Humidity at 925 hPa	9 / 153
<b>T1000</b>	Temperature at 1000 hPa	5 / 85
<b>T2</b>	Temperature at 2 hPa	5 / 51
<b>T200</b>	Temperature at 200 hPa	5 / 61
<b>T50</b>	Temperature at 50 hPa	5 / 52
<b>T500</b>	Temperature at 500 hPa	5 / 68
<b>T700</b>	Temperature at 700 hPa	5 / 79
<b>T850</b>	Temperature at 850 hPa	5 / 86
<b>VVEL500</b>	Vertical Velocity at 500 hPa	12 / 306
<b>WND200</b>	U and V Wind at 200 hPa	15 / 202
<b>WND500</b>	U and V Wind at 500 hPa	20 / 360
<b>WND700</b>	U and V Wind at 700 hPa	20 / 393
<b>WND850</b>	U and V Wind at 850 hPa	20 / 400
<b>WND1000</b>	U and V Wind at 1000 hPa	20 / 384
<b>Z200</b>	Geopotential at 200 hPa	13 / 161
<b>Z500</b>	Geopotential at 500 hPa	13 / 175
<b>Z700</b>	Geopotential at 700 hPa	16 / 200
<b>Z850</b>	Geopotential at 850 hPa	17 / 206
<b>Z1000</b>	Geopotential at 1000 hPa	18 / 221

### **Hourly Timeseries from OCN files (13 parameters)**

*(analysis with 0-6 hour forecast)*

Low resolution 1 degree / High resolution 0.5 degree

<b><i>File</i></b>	<b><i>Description</i></b>	<b><i>(low/high res in MB/month)</i></b>
<b>OCNDT20C</b>	Depth of 20C Isotherm	15 / 50
<b>OCNHEAT</b>	Ocean Heat Content	55 / 155
<b>OCNMLD</b>	Ocean Mixed Layer Depth	27 / 95
<b>OCNSAL5</b>	Ocean Salinity at depth of 5m	23 / 75
<b>OCNSAL15</b>	Ocean Salinity at depth of 15m	23 / 75
<b>OCNSLH</b>	Sea Level Height	34 / 118
<b>OCNSST</b>	Ocean Potential Temperature at depth of 5m	34 / 118
<b>OCNT15</b>	Ocean Potential Temperature at depth of 15m	34 / 118
<b>OCNU5</b>	Ocean Zonal Current at depth of 5m	36 / 130
<b>OCNV5</b>	Ocean Meridional Current at depth of 5m	36 / 130
<b>OCNU15</b>	Ocean Zonal Current at depth of 15m	36 / 130
<b>OCNV15</b>	Ocean Meridional Current at depth of 15m	36 / 130
<b>OCNVV55</b>	Ocean vertical velocity at depth of 55 m	72 / 212

### **Hourly Timeseries from IPV files (3 parameters)**

*(analysis with 0-6 hour forecast)*

Low resolution 2.5 degree / High resolution 0.5 degree

<b><i>File</i></b>	<b><i>Description</i></b>	<b><i>(low/high res in MB/month)</i></b>
<b>IPV450</b>	Potential Vorticity at 450 K Isentropic	11 / 140
<b>IPV550</b>	Potential Vorticity at 550 K Isentropic	11 / 140
<b>IPV650</b>	Potential Vorticity at 650 K Isentropic	11 / 140

### **Hourly Output Archives**

<b><i>Type</i></b>	<b><i>Description</i></b>	<b><i>Output Files</i></b>	<b><i>Per file MB</i></b>
<b>FLXF</b>	104 records, sfc, radiative flux vars	f00-f06, f09	30
<b>PGBH</b>	628 records, pressure level vars	anl, f00-f06, f09	90
<b>OCNH</b>	222 records, ocean variables	f01-f06, f09	35
<b>DIABF</b>	926 records, diabatic heating, etc.	f00-f06, f09	26
<b>IPVH</b>	130 records, isentropic level	anl, f00-f06, f09	17

### **Initial Conditions Archive (every 6 hours)**

Low and high resolution approximate file size

<b><i>File</i></b>	<b><i>Description</i></b>	<b><i>(low/high res MB)</i></b>
<b>SIGANL</b>	3-D Hybrid Analysis (Binary)	230 / 25
<b>SFCANL</b>	Surface Analysis (Binary)	120 / 13
<b>OCNANL</b>	Ocean Analysis (Binary)	1250 ( <i>high res only</i> )

## **Appendix C: *The Data Access***

To address a growing need for remote access to high volume numerical weather prediction and global climate models and data, NOAA's National Climatic Data Center (NCDC), along with the National Centers for Environmental Prediction (NCEP) and the Geophysical Fluid Dynamics Laboratory (GFDL), initiated the NOAA Operational Model Archive and Distribution System (NOMADS, Rutledge, et al., 2006) project. The NOMADS framework was developed to facilitate climate model and observational data inter-comparison capabilities as discussed in documents such as the Intergovernmental Panel on Climate Change (IPCC, 2001), and advances a direct recommendation by the National Academies of Science, Board on Atmospheric Sciences and Climate (BASC) to improve multi-model ensemble diagnostics capabilities (National Academies of Science, NRC, 2007).

Of the 250Tb CFSR total output (estimated at the time of this writing), approximately 10Tb are considered to be the "Tier-1" or "high priority" data. These data were determined by several requirements workshops conducted by NCEP, and at an NCDC hosted Reanalysis Town Hall Workshop at the January 2008 Annual Meeting of the American Meteorological Society. At the AMS meeting NCDC obtained feedback on user requirements for this massive dataset. An outcome of that Workshop culminated in the generation of a user web page soliciting additional user requirements. The NOAA Reanalysis Community Web Forum can be accessed at: <http://nomads.ncdc.noaa.gov/NOAAREanalysis/cfsrr>.

The primary NOAA access point for CFSR is NCDC NOMADS located at: <http://nomads.ncdc.noaa.gov>. On NOMADS data access to CFSR is facilitated by file and variable type organization most requested by users based on previous reanalysis outputs. The "Tier-1" data include a suite of variables aggregated in a long time-series, and a suite of monthly mean variables. The monthly mean and time series data are available as separate files by variable type, month and year for the period of record to facilitate user access. A separate dataset of monthly mean data have been aggregated for access as individual files of specific type so users can make a single request for the entire period of record of individual variables (e.g., 31 years of 500hPa heights). These "long time-series" datasets will greatly reduce processing time on the NCDC NOMADS while also providing improved user access response times.

To facilitate user access the National Center for Atmospheric Research (NCAR) has agreed to act as a mirror site to the NCDC's NOMADS access point and is scheduled to provide access to CFSR. The NCAR access point is: <http://dss.ucar.edu/>.

Given the massive volume of these data, users are cautioned that extremely high volume data requests may be throttled at NCDC based on the size of the individual request and the number of concurrent users.



## **Appendix D: Details of Satellite bias correction spin up experiments for CFSR**

This CFSR covers the period from 1979 to 2008. During these 30 years, a series of NOAA polar orbiting satellites operated for periods of these years (Figure 21). The radiance measurements from these satellites were directly assimilated into the analysis system, which is one of the major improvements of the CFSR over R2. Substantial biases exist when the radiance measurements are compared to the model-simulated radiances. These biases are complicated and relate to instrument calibration, data processing and deficiencies in the radiative transfer model. Accurate bias corrections for these satellite observations were required for the success of the CFSR. This part of the paper describes the method to estimate the biases of these historical satellite observations.

A variational satellite bias correction scheme was introduced by Derber and Wu (1998) when direct assimilation of radiance began. This scheme has been continually developed at NCEP since then and is used in the GSI system adapted for the CFSR. Radiance bias is composed of two parts: the satellite scan angle dependent bias (SATANG) and the air-mass dependent part (BIASCR). The SATANG file is updated following the analysis step as the weighted average of the previous cycle's SATANG file (containing the bias) and the departure between the new radiance measurements and the model-simulated radiances, o-g, (where o is the observation and g is the forecast guess). SATANG is allowed to evolve very slowly (the weight given to the new (o-g) is 1/120). The air mass bias is predicted as the linear combination of five predictors and their corresponding coefficients, the BIASCR file. The five predictors are, respectively, a constant 0.01, the solar zenith angle, the cloud liquid water, CLW, (only applied to microwave instruments), the temperature lapse rate and the square of the lapse rate. The predictor coefficients are calculated using the variational method (the predictor coefficients are additional analysis variables and included in the minimization).

The biases from the historical satellite radiances can be estimated using a two-step procedure. First, the GDAS is run for a training period, starting from guess values, and the statistics of the (o-g) are collected throughout this training period. Second, the collected radiance statistics are then post-processed to calculate a SATANG file. For a given scan position, channel, sensor, and satellite, the SATANG value is the average of the (o-g) weighted by the specified observation error inverse over the whole globe and the entire training period.

Several preliminary experiments have been carried out to determine starting values and how long the training period should be. These experiments were run over the period of 2006110118 to 2007013018 using zeroed out SATANG and BIASCR files. To simulate historical MSU and HIRS2 channels 2-14, only AMSUA-N15 channels 3, 5, 7 and 9 and HIRS3-N17 channels 2-14, respectively, were assimilated. (The CRTM SSU cell pressure correction noted in section 3.1.2 was still in development while these tests were run.) The weight given to the new (o-g) is 0.033 which is larger than 0.00833 used in the analysis, since the purpose was to spin up the bias correction. The preliminary experiments show that the predictor coefficients, except for the CLW term, can start from zero and spin up rapidly, usually within a month (not shown). The coefficients for the CLW evolve very slowly, as shown in Figure 25, thus, the CLW coefficients from NOAA-15 are used as guess values of the training period for all the historical MSU channels.

Figure 26 shows the global total bias (air-mass part plus scan angle part) for the active channels of the AMSUA-N15 and HIRS3-N17. In this figure, CNTL is the operational GDAS total bias for these channels as a reference. Most of the channels (o-g) equilibrate within about one month, except for HIRS3\_N17 channel 12, which took about 2.5 months. Therefore, 3 months was chosen as the general training length for all the historical satellites.

The next prerequisite to running the spin-up experiments was to assemble a time series of active channels and sensors for all the historical satellites, the SATINFO files, marked with periods of known outages and poor quality data. The starting point was a set of tables and scripts received, via JCSDA collaboration, from the MERRA reanalysis, see Bosilovich, 2008. The tables and scripts were updated based upon the ERA-40 quality control list and the NCEP historical satellite document maintained by co-author Dennis Keyser, at:

[http://www.emc.ncep.noaa.gov/mmb/data\\_processing/Satellite\\_Historical\\_Documentation.htm](http://www.emc.ncep.noaa.gov/mmb/data_processing/Satellite_Historical_Documentation.htm).

Once a working set of SATINFO files were created (and the historical set of SSU instruments were included in the CRTM as noted in 3.1.2), a set of bias correction spin up experiments listed in Table 3 was carried out by running the CFSR over the indicated periods when the satellite data first became available (See Figure 21). In 2 cases, NOAA-6 and 11 operated during two periods, so the bias correction spin up had to be done twice. Once each spin up experiment was complete, the post-processing step was carried out on the (o-g) diagnostic files to create starting SATANG files. It was then paired with the BIASCR file from the last cycle of the training period for use when the final CFSR assimilation began assimilating that particular instrument and sensor. Two features of this bias correction scheme should be noted. One is the scan angle dependent bias is the dominant part of the total bias, and the other is the predictor coefficients of the air-mass bias usually responds to the atmospheric state very quickly, usually within one or two days.

## REFERENCES

- Accadia, C., S. Mariani, M. Casaioli, and A. Lavagnini, 2003: Sensitivity of Precipitation Forecast Skill Scores to Bilinear Interpolation and a Simple Nearest-Neighbor Average Method on High-Resolution Verification Grids, *Wea. Forecasting*, **18**, 918-932.
- Akmaev, R.A., and H.-M.H. Juang, 2008: Using enthalpy as a prognostic variable in atmospheric modeling with variable composition. *Quart. J. Roy. Meteor. Soc.*, **134**, 2193-2197.
- Alpert, J. C., M. Kanamitsu, P. M. Caplan, J.G. Sela, G. H. White, and E. Kalnay, 1988: Mountain induced gravity wave drag parameterization in the NMC medium-range forecast model. *Proc. 8<sup>th</sup> Conf. on NWP*, Baltimore, MD.
- Alpert, J. C., S.-Y. Hong, and Y.-J. Kim, 1996: Sensitivity of cyclogenesis to lower tropospheric enhancement of gravity wave drag using the environmental modeling center medium range model. *Proc. 11<sup>th</sup> Conf. on NWP*, Norfolk, VA.
- Alpert, J.C., 2004: Subgrid-scale Mountain blocking at NCEP. *Proc. 20<sup>th</sup> Conf. on Weather and Forecasting*, Seattle, WA.
- Andersson, E., and H. Järvinen, 1999: Variational quality control. *Quart. J. R. Meteor. Soc.*, **125**, 697-722.
- Anderson, J., T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, 2009: The Data Assimilation Research Testbed: A Community Facility. *Bull. Amer. Meteor. Soc.*, **90**, 1283–1296.

- Argo Science Team, 2001: The global array of profiling floats, *Observing the Ocean in the 21<sup>st</sup> Century*, C.J. Koblinsky and N.R. Smith, Eds., Australian Bureau of Meteorology, 248-258.
- Assel, Raymond A., David C. Norton, and Kevin C. Cronk, 2002: A Great Lakes Ice Cover Digital Data Set for Winters 1973-2000, *NOAA-Technical Memorandum GLERL-121*, 46 pp.
- Balaji, V., 2007: The FMS coupler architecture, Ocean Model Coupling Workshop, Princeton, NJ, Feb 7, 2007.
- Behringer, D. W., Ji, M., and A. Leetmaa, 1998: An improved coupled model for ENSO prediction and implications for ocean initialization. Part I: The ocean data assimilation system, *Mon. Wea. Rev.*, **126**, 1013-1021.
- Behringer, D. W., and Y. Xue, 2004: Evaluation of the global ocean data assimilation system at NCEP: The Pacific Ocean, *Eighth Symposium on Integrated Observing and Assimilation System for Atmosphere, Ocean, and Land Surface, AMS 84<sup>th</sup> Annual Meeting*, Seattle, Washington, 11-15.
- Behringer, D. W., 2007: The Global Ocean Data Assimilation System at NCEP, *11<sup>th</sup> Symposium on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface, AMS 87th Annual Meeting*, San Antonio, Texas, 12pp
- Bender, M.A., I. Ginis, R. Tuleya, B. Thomas, and T. Marchok, 2007: The operational GFDL coupled hurricane-ocean prediction system and summary of its performance. *Mon. Wea. Rev.*, **135**, 3965-3989.
- Bengtsson, L., P. Arkin, P. Berrisford, P. Bougeault, C.K. Folland, C. Gordon, K. Haines, K.I. Hodges, P. Jones, P. Kallberg, N. Rayner, A.J. Simmons, D. Stammer, P.W. Thorne, S. Uppala, and R.S. Vose, 2007: The Need for a Dynamical Climate Reanalysis. *Bull. Amer. Meteor. Soc.*, **88**, 495-501.
- Bosilovich, Michael, 2008. NASA's Modern Era Retrospective-analysis for Research and Applications: Integrating Earth Observations. Earthzine. E-Zine Article.

- Bourlès, B., R. Lumpkin, M. J. McPhaden, F. Hernandez, P. Nobre, E. Campos, L. Yu, S. Planton, A. J. Busalacchi, A. D. Moura, J. Servain, and J. Trotte, 2008: The PIRATA Program: History, Accomplishments, and Future Directions. *Bull. Amer. Meteor. Soc.*, **89**, 1111-1125.
- Cavalieri, D. J., 1994: Sea ice algorithm in NASA Sea Ice Validation Program for the Defense Meteorological Satellite Program Special Sensor Microwave Imager: Final Report, *NASA Technical Memorandum* 104559, pp. 25-32, 1992.
- Cavalieri, D., C. Parkinson, P. Gloersen, and H. J. Zwally, 2007: Sea ice concentrations from Nimbus-7 SMMR and DMSP SSM/I passive microwave data, 1978-1996. Boulder, Colorado USA: National Snow and Ice Data Center. Digital media. 1996, updated 2007.
- Chou, M.D, M. J. Suarez, C. H. Ho, M. M. H. Yan, and K. T. Lee, 1998: Parameterizations of cloud overlapping and shortwave single scattering properties for use in general circulation and cloud ensemble models. *J. Climate*, **11**, 202-214.
- Clough, S.A., M.W. Shephard, E.J. Mlawer, J.S. Delamere, M.J. Iacono, K. Cady-Pereira, S. Boukabara, and P.D. Brown, 2005: Atmospheric radiative transfer modeling: a summary of the AER codes, *J. Quant., Spectrosc. Radiat. Transfer*, **91**, 233-244.
- Collins N., G. Theurich, C. DeLuca, M. Suarez, A. Traynaov, V. Balaji, P. Li, W. Yang, C. Hill, and A. da Silva, 2005: Design and implementation of components of Earth System modeling Framework. *The International Journal of High Performance Computing Applications*, **19**, #3, Summer 2005, pp-355-356.
- Collins, W.G. and L.S. Gandin, 1990: Comprehensive hydrostatic quality control at the National Meteorological Center. *Mon. Wea. Rev.*, **18**, 2754-2767.

- Compo, G.P., J.S. Whitaker, and P.D. Sardeshmukh, 2006: Feasibility of a 100-Year Reanalysis Using Only Surface Pressure Data. *Bull. Amer. Meteor. Soc.*, **87**, 175–190.
- Conkright, M. E., S. Levitus, T. O'Brien, T. P. Boyer, C. Stephens, D. Johnson, O. Baranova, J. Antonov, R. Gelfeld, J. Rochester, and C. Forgy, 1999: *World Ocean Database 1998, Documentation and Quality Control Version 2.0*, National Oceanographic Data Center Internal Report 14, National Oceanographic Data Center, Silver Spring, Maryland.
- Cucurull, L. and J. C. Derber, 2008: Operational implementation of COSMIC observations into the NCEP's Global Data Assimilation System, *Wea. And Forecasting*, **23**, 4, 702-711.
- Cucurull, L., 2009: Improvement in the use of an operational constellation of GPS Radio-occultation receivers in Weather Forecasting, *Wea. and Forecasting*, in press.
- Dai, A., and J. Wang, 1999: Diurnal and Semidiurnal Tides in Global Surface Pressure Fields. *J. Atmos. Sci.*, **56**, 3874–3891.
- Davis, G., 2007: "History of the NOAA satellite program", *Journal of Applied Remote Sensing*, **1**, 012504.
- Deardorff, J.W., 1980: Cloud top entrainment instability, *J. Atmos. Sci.*, **37**, 131-147
- Derber, J., and A. Rosati, 1989: A global oceanic data assimilation system, *J. Phys. Oceanogr.*, **19**, 1333-1347
- Derber, J.C., D. F. Parrish, and S. J. Lord, 1991: The new global operational analysis system at the National Meteorological Center. *Wea. Forecasting*, **6**, 538-547.
- Derber, J. C. and W.-S. Wu, 1998: The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Wea. Rev.*, **126**, 2287 - 2299.

- de Viron, O., G. Schwarzbaum, F. Lott, and V. Dehant, 2005: Diurnal and subdiurnal effects of the atmosphere on the Earth rotation and geocenter motion, *J. Geophys. Res.*, **110**, B11404, doi:10.1029/2005JB003761.
- Dworak, R. and J. Key, 2007: 20 Years of Polar Winds from AVHRR: Validation and Comparison to the ERA-40. *J. Atmos. Ocean. Tech.*, accepted (May 2008).
- Ek M.B., K.E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J.D. Tarplay, 2003: Implementation of the Noah land-use model advances in the NCEP operational mesoscale Eta model. *J. Geophys. Res.*, **108**, 8851, doi:10.1029/2002JD003296.
- English, S.J., R.J. Renshaw, P.C. Dibben, A.J. Smith, P.J. Rayer, C. Poulsen, F.W. Saunders, and J.E. Eyre, 2000: A comparison of the impact of TOVS and ATOVS satellite sounding data on the accuracy of numerical weather forecasts. *Quart. J. R. Met. Soc.*, **126**, 2911-2931
- Fels, S. and M. D. Schwarzkopf, 1975: The simplified exchange approximation: a new method for radiative transfer calculations. *J. Atmos. Sci.*, **32**(7), 1475-1488.
- Fiorino, M. 2002: Analysis and forecasts of tropical cyclones in the ECMWF 40-year reanalysis (ERA-40). *ERA-40 Project Report Series*, **3**. Workshop on Re-analysis, 5-9 Nov 2001. ECMWF, Shinfield Park, UK, 443 pp.
- Fiorino, M. 2004: A Multi-decadal Daily Sea Surface Temperature & Sea Ice Concentration Data Set for the ECMWF ERA-40 Reanalysis. *ERA-40 Project Report Series* No. 12, European Centre for Medium-Range Weather Forecasts, Reading, UK, 16 pp  
[http://www.ecmwf.int/publications/library/ecpublications/pdf/ERA40\\_PRS12.pdf](http://www.ecmwf.int/publications/library/ecpublications/pdf/ERA40_PRS12.pdf)
- Flynn, L., D. McNamara, C.T. Beck, I. Petropavlovskikh, E. Beach, Y. Pachepsly, Y.P. Li, M., Deland, L.-K. Huang, C. Long, R. Tiruchirapalli, and S. Taylor, 2009: Measurements and products from the Solar Backscatter Ultraviolet (SBUV/2) and

- ozone Mapping and Profiler Suite (OMPS) instruments, *Intl. J. Rem. Sens.*, **30**, No. 15, 4259-4272, 2009.
- Fu, X., and B. Wang, 2004: Differences of Boreal Summer Intraseasonal Oscillations simulated in an Atmosphere-Ocean Coupled Model and an Atmosphere-Only Model. *J. Clim.*, **17**, 1263-1271.
- Gandin, L.S.: 1988: Complex quality control of meteorological observations. *Mon. Wea. Rev.*, **116**, 1138-1156.
- Gaspari, G., S.E. Cohn, J. Guo and S. Pawson, 2006: Construction and Application of Covariance Functions with Variable Length Fields. *Q. J. Roy. Meteorol. Soc.*, **132**, Part B, No. 619, 1815-1838.
- Gemmill, W. H. and V. M. Krasnopolsky, 1999: The Use of SSM/I Data in Operational Marine Analysis, *Wea. and Forecasting*, **14**, 789-800.
- Gemmill, W., B. Katz, and X. Li, 2007: Daily Real-Time Global Sea Surface Temperature, A High-Resolution Analysis: RTG\_SST\_RH, MMAB Technical Note 260, 39 pp.
- Gent, P. R., and J. C. McWilliams, 1990: Isopycnal mixing in ocean circulation models. *Journal of Physical Oceanography*, **20**, 150–155.
- Gibson, J.K., P. Kallberg, S. Uppala, A. Hernandez, A. Nomura and E. Serrano, 1997: ERA Description. *ECMWF Re-Analysis Project Report Series*, 1.
- Griffies, S. M., A. Gnanadesikan, R. C. Pacanowski, V. Larichev, J. K. Dukowicz, and R. D. Smith, 1998: Isonutral diffusion in a  $z$ -coordinate ocean model. *Journal of Physical Oceanography*, **28**, 805–830.
- Griffies, S. M., A. Gnanadesikan, R. C. Pacanowski, V. Larichev, J. K. Dukowicz, and R. D. Smith, 1998: Isonutral diffusion in a  $z$ -coordinate ocean model. *Journal of Physical Oceanography*, **28**, 805–830.



- Griffies, S. M., and R. W. Hallberg, 2000: Biharmonic friction with a Smagorinsky viscosity for use in large-scale eddy-permitting ocean models. *Monthly Weather Review*, **128**, 2935–2946
- Griffies, S.M., M.J. Harrison, R.C. Pacanowski, and A. Rosati, 2004: Technical Guide to MOM4, GFDL Ocean Group Technical Report No. 5, NOAA/Geophysical Fluid Dynamics Laboratory. Available on-line at <http://www.gfdl.noaa.gov/~fms>
- Grumbine, R. W., 1996: Automated Passive Microwave Sea Ice Concentration Analysis at NCEP, OMB Tech. Note 120, NCEP, 13 pp.
- Grumbine, R. W., 2009: *A posteriori* filtering of sea ice concentration analyses, MMAB Tech. Note 282, NCEP, 7 pp
- Grumbine, R. W., 2010a: Operational NCEP sea ice concentration analysis. In preparation.
- Grumbine, R. W., 2010b: CFSR Sea ice analysis. In preparation.
- Hamill, T.M., J.S. Whitaker, and S.L. Mullen, 2006: Reforecasts: An Important Dataset for Improving Weather Predictions. *Bull. Amer. Meteor. Soc.*, **87**, 33–46.
- Han, Y., P. van Delst, Q. Liu, F. Weng, B. Yan, R. Treadon, and J. Derber, 2006: JCSDA Community Radiative Transfer Model (CRTM) - Version 1. NOAA Tech. Rep. NESDIS 122, 33 pp.
- Haurwitz, B. and D. Cowley, 1973: The diurnal and semidiurnal barometric oscillations, global distribution and annual variation. *Pure Appl. Geophys.*, **102**, 193–222..
- Helfrich, S.R., D. McNamara, B.H. Ramsay, T. Baldwin and T. Kasheta, 2007: Enhancements to, and forthcoming developments in the Interactive Multisensor Snow and Ice Mapping System (IMS). *Hydrol. Process.*, **21**, 1576–1586, doi:10.1002/hyp.6720.
- Hong, S. -Y. and H. -L. Pan, 1996: Nonlocal boundary layer vertical diffusion in a

- medium range forecast model. *Mon. Wea. Rev.*, **124**, 2322-2339.
- Hong, S. -Y. and H. -L. Pan, 1998: Convective Trigger Function for a Mass-Flux Cumulus Parameterization Scheme. *Mon. Wea. Rev.*, **126**, 2599–2620.
- Hou, Y. -T., K. A. Campana and S.–K. Yang, 1996: Shortwave radiation calculations in the NCEP's global model. *Int. Radiation Symposium*, IRS-96, August 19-24, Fairbanks, AL.
- Hou, Y., S. Moorthi, and K. Campana, 2002: Parameterization of solar radiation transfer in the NCEP models. *NCEP Office Note*, #441. [Available online at: <http://www.emc.noaa.gov/officenotes/FullTOC.html#2000> ].
- Iacono, M.J., E.J. Mlawer, S.A. Clough, and J.-J. Morcrette, 2000: Impact of an improved longwave radiation model, RRTM, on the energy budget and thermodynamic properties of the NCAR Community Climate Model, CCM3, *J. Geophys. Res.*, **105**, 14873-14890, 2000.
- Jenne, R., and J. Woollen, 1994, The Reanalysis Database, Extended Abstracts, Tenth Conference on Numerical Weather Prediction, Portland, OR, American Meteorological Society.
- Ji, M., A. Leetmaa, and J. Derber, 1995: An ocean analysis system for seasonal to interannual climate studies, *Mon. Wea. Rev.*, **123**, 460-481.
- Juang, H.-M., 2005: Discrete generalized hybrid vertical coordinates by a mass, energy and angular momentum conserving vertical finite-differencing scheme. NCEP Office Note 445, 33 pp. [Available online at: <http://www.emc.ncep.noaa.gov/officenotes/FullTOC.html#2000> ].
- Juang, H.-M.H, 2009: A multi-conserving discretization with enthalpy as thermodynamic prognostic variable in generalized hybrid vertical coordinates for the NCEP

- Global Forecast System. Manuscript submitted to *Mon. Wea. Rev.*
- Kadi, M., 2009: The AMMA Observing Network Contribution to GCOS, GCOS STEERING COMMITTEE, 17<sup>TH</sup> Session, Paris, France, 27-30 October 2009
- Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, A. Leetmaa, B. Reynolds, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. Mo, C. Ropelewski, J. Wang, R. Jenne, and D. Joseph, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471.
- Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.K. Yang, J.J. Hnilo, M. Fiorino, and G.L. Potter, 2002: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, **83**, 1631–1643.
- Kistler, R., E. Kalnay, W. Collins, S. Saha, G. White, J. Woollen, M. Chelliah, W. Ebisuzaki, M. Kanamitsu, V. Kousky, H. van den Dool, R. Jenne, and M. Fiorino, 2001: The NCEP–NCAR 50–Year Reanalysis: Monthly Means CD–ROM and Documentation. *Bull. Amer. Meteor. Soc.*, **82**, 247–267.
- Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, R. M. Errico, and R. Yang, 2009: Improving Incremental Balance in the GSI 3DVAR Analysis System, *Mon. Wea. Rev.*, **137**, 1046-1060.
- Kobayashi, S., M. Matricardi, D. Dee, and S. Uppala 2009: Toward a consistent reanalysis of the upper stratosphere based on radiance measurements from SSU and AMSU-A. *Quart. J. of the Royal Meteor. Soc.*, **135**, 2086 - 2099
- Kopp, T. J. and R.B. Kiess, 1996: The Air Force Global Weather Central snow analysis model. Preprints, 15<sup>th</sup> Conf. on Weather Analysis and Forecasting, Norfolk, VA, *Amer. Meteor. Soc.*, 220-222.
- Krasnopolsky, V.M., L.C. Breaker and W.H. Gemmill, 1995: A neural network as a nonlinear transfer function model for retrieving surface wind speeds from the

- SpecialSensor Microwave/Imager. *J. Geophys. Res.*, **100**, No. C6, 11033-11045.
- Large, W. G., J. C. McWilliams and S. C. Doney, 1994: Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. *Reviews of Geophysics*, **32**, 363–403.
- Lindzen, R. S., and J.R. Holton, 1968: A theory of quasi-biennial oscillation. *J. Atmos. Sci.*, **26**, 1095-1107.
- Lott, F. and M. J. Miller, 1997: A new subgrid-scale orographic drag parameterization: its performance and testing. *Quart. J. Roy. Meteor. Soc.*, **123**, 101-127.
- Liu, Q., T. Marchok, H. Pan, M. Bender and S. Lord, 1999: Improvements in hurricane initialization and forecasting at NCEP with global and regional (GFDL) models. NOAA Technical Procedures Bulletin No. 472, National Weather Service, Office of Meteorology, 1325 East West Highway, Silver Spring MD, 20910  
<http://www.nws.noaa.gov/om/tpb/472.htm>
- Liu, Q. and F. Weng, 2006, Advanced Doubling–Adding Method for Radiative Transfer in Planetary Atmospheres, *J. Atmos. Sci.* **63**, 3459-3465.
- Liu, Q., and F. Weng, 2009: Recent stratospheric temperature observed from satellite measurements, *SOLA*, **5**, 53-56, doi:10.2151/sola.2009-014.
- Lumpkin, R. and M. Pazos, 2006: Measuring surface currents with Surface Velocity Program drifters: the instrument, its data, and some recent results. Chapter two of Lagrangian Analysis and Prediction of Coastal and Ocean Dynamics (LAPCOD) ed. A. Griffa, A. D. Kirwan, A. J. Mariano, T. Ozgokmen, and T. Rossby.
- Markus, T. and D. J. Cavalieri, 2000: An enhancement of the NASA Team sea ice algorithm, *IEEE Trans. Geoscience Remote Sensing*, **38**, 1387-1398.
- Markus, T. and D. J. Cavalieri, 2009: The AMSR-E NT2 sea ice concentration algorithm: its basis and implementation, *J. of the Remote Sensing Soc of Japan*, **29**, 1, 216-225.
- McCormack, J.P., S.D., Eckermann, D.E. Siskind and T. McGee, 2006: CHEM2D-OPP: A new linearized gas phase photochemistry parameterization for high altitude NWP

- and climate models. *Atmos. Chem. Phys.*, **6**, 4943-4972.
- McNally, A.P., J.C. Derber, W.-S. Wu and B.B. Katz, 2000: The use of TOVS level-1B radiances in the NCEP SSI analysis system. *Quart.J.R.Meteor.Soc.*, **126**, 689-724.
- McPhaden, M. J., A. J. Busalacchi, R. Cheney, J. R. Donguy, K. S. Gage, D. Halpern, M. Ji, P. Julian, G. Meyers, G. T. Mitchum, P. P. Niiler, J. Picaut, R. W. Reynolds, N. Smith and K. Takeuchi, 1998: The Tropical Ocean-Global Atmosphere (TOGA) observing system: A decade of progress. *J. Geophys. Res.*, **103**, 14,169-14,240.
- McPhaden, M. J., G. Meyers, K. Ando, Y. Masumoto, V. S. N. Murty, M. Ravichandran, F. Syamsudin, J. Vialard, L. Yu, and W. Yu, 2009: RAMA: The Research Moored Array for African–Asian–Australian Monsoon Analysis and Prediction. *Bull. Amer. Meteor. Soc.*, **90**, 459 - 480.
- Mesinger, F., G. DiMego, E. Kalnay, K. Mitchell, P.C. Shafran, W. Ebisuzaki, D. Jović, J. Woollen, E. Rogers, E.H. Berbery, M.B. Ek, Y. Fan, R. Grumbine, W. Higgins, H. Li, Y. Lin, G. Manikin, D. Parrish, and W. Shi, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360.
- Mlawer E. J., S. J. Taubman, P. D. Brown, M.J. Iacono and S.A. Clough, 1997: radiative transfer for inhomogeneous atmosphere: RRTM, a validated correlated-K model for the longwave. *J. Geophys. Res.*, **102**(D14), 16, 663-16, 6832.
- Mo. T., M. D. Goldberg, D. S. Crosby, and Z. Cheng, 2001: Recalibration of the NOAA microwave sounding unit. *J. Geophys. Res.*, **106**, 10145–10150.
- Moorthi, S., H. L. Pan and P. Caplan, 2001: Changes to the 2001 NCEP operational MRF/AVN global analysis/forecast system. *NWS Technical Procedures Bulletin*, **484**, pp14. Available at: <http://www.nws.noaa.gov/om/tpb/484.htm>
- Moorthi, S., R. Sun, H. Xiao, and C. R. Mechoso, 2010: Low-cloud simulation in the Southeast Pacific in the NCEP GFS: role of vertical mixing and shallow convection. *NCEP Office Note # 463*. Available at: <http://www.emc.ncep.noaa.gov/officenotes/FullTOC.html#2000>

- Murray, R. J., 1996: Explicit generation of orthogonal grids for ocean models. *Journal of Computational Physics*, **126**, 251–273.
- National Academies of Science, National Research Council (NRC), Board on Atmospheric Sciences and Climate (BADC), 2007: Completing the Forecast: Characterizing and Communicating Uncertainty for Better Decisions Using Weather and Climate Forecasts. pp76.
- Neumann, C. J. (1999), The HURISK model: An adaptation for the Southern Hemisphere (A user's manual) report, contract N00014-96-C-6015, 31 pp., Sci. Appl. Int. Corp., Monterey, Calif.
- Onogi, K., J. Tsutsui, H. Koide, M. Sakamoto, S. Kobayashi, H. Hatsushika, T. Matsumoto, N. Yamazaki, H. Kamahori, K. Takahashi, S. Kadokura, K. Wada, K. Kato, R. Oyama, T. Ose, N. Mannoji and R. Taira, 2007: The JRA-25 Reanalysis. *J. Meteor. Soc. Japan*, **85**, 3, 369-432.
- Pan, H. L., and L. Mahrt, 1987: Interaction between soil hydrology and boundary layer development, *Boundary Layer Meteorology*, **38**, 185-202.
- Pan, H.-L. and W.-S. Wu, 1995: Implementing a mass flux convective parameterization package for the NMC medium range forecast model. NMC Office Note 409, 40 pp.[Available online at:  
<http://www.emc.ncep.noaa.gov/officenotes/FullTOC.html#1990> ]
- Paolino, D., Q. Yang, B. Doty, J. Kinter, J. Shukla, and D.M. Straus, 1995: A Pilot Reanalysis Project at COLA. *Bull. Amer. Meteor. Soc.*, **76**, 697–710.
- Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical interpolation system. *Mon. Wea. Rev.*, **120**, 1747-1763.
- Pegion K and B. Kirtman 2008: The Impact of Air-Sea Interactions on the Simulation of Tropical Intraseasonal Variability. *J. Climate* **24**, 6616-6635.
- Peters-Lidard , C.D., P.R. Houser, Y. Tian, S.V. Kumar, J. Geiger, S. Olden, L. Lighty, B. Doty, P. Dirmeyer, J. Adams, K. Mitchell, E.F. Wood and J. Sheffield, 2007: *High-*

*performance Earth system modeling with NASA/GSFC's Land Information System.*  
Innovations in Systems and Software Engineering. **Vol.3(3)**, 157-165.  
doi:10.1007/s11334-007-0028-x.

- Phillips, N. A., 1957: A coordinate system with some special advantages for numerical forecasting *J. Atmos. Sci.*, **14**, 184–185.
- Ponte, R. M., and R. D. Ray, 2002: Atmospheric pressure corrections in geodesy and oceanography: A strategy for handling air tides, *Geophys. Res. Lett.*, **29**(24), 2153, doi:10.1029/2002GL016340.
- Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts, 2003a: Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian covariances. *Mon. Wea. Rev.*, **131**, 1524-1535.
- Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts, 2003b: Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially inhomogeneous and anisotropic general covariances. *Mon. Wea. Rev.*, **131**, 1536-1548.
- Rančić M., J. C. Derber, D. Parrish, R. Treadon, and D. T. Kleist, 2008: The development of the first-order time extrapolation to the observation (FOTO) method and its application in the NCEP global data assimilation system. *Proc. 12th Symp. on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS)*, New Orleans, LA, Amer. Meteor. Soc., CD-ROM, J6.1.
- Randall, D. A., 1980: Conditional instability of the first kind upside-down. *J. Atmos. Sci.*, **37**, 125-130.
- Ray, R. D., and R. M. Ponte, 2003: Barometric tides from ECMWF operational analyses, *Ann. Geophys.*, **21**, 1897–1910.

- Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax, 2007: Daily high-resolution blended analyses for sea surface temperature. *J. Climate*, **20**, 5473-5496.
- Rocken, C., R., and co-authors, 1997: Analysis and validation of GPS/MET data in the neutral atmosphere. *J. Geophys. Res.*, **102** (D25), 29849-29866.
- Rutledge, G.K., J. Alpert, and W. Ebuisaki, 2006: NOMADS: A Climate and Weather Model Archive at the National Oceanic and Atmospheric Administration. *Bull. Amer. Meteor. Soc.*, **87**, 327-341.
- Saha, S., S. Nadiga, C. Thiaw, J. Wang, W. Wang, Q. Zhang, H. M. van den Dool, H.L. Pan, S. Moorthi, D. Behringer, D. Stokes, M. Peña, S. Lord, G. White, W. Ebisuzaki, P. Peng, and P. Xie, 2006: The NCEP Climate Forecast System. *J. Climate*, **19**, 3483–3517.
- Schubert, S. D., W. Min, L. Takacs and J. Joiner, 1997: Reanalysis of historical observations and its role in the development of the Goddard EOS Climate Data Assimilation System. *Advances in Space Research*, **19**, Issue 3, 1997, Pages 491-501
- Schwarzkopf, M.D., and S. Fels, 1991: The simplified exchange method revisited: An accurate, rapid method for computation of infrared cooling rates and fluxes. *J. Geophys Res.*, **96(D5)**, 9075-9096.
- Seaman, R. and Hart, T., 2003: The history of PAOBs in the Australian Bureau of Meteorology, *Aust. Met. Mag.* **52** (2003) 241-250.
- Smith, W. L., H. M. Woolf, C. M. Hayden, D. Q. Wark, and L. M. McMillin, 1979: The TIROS-N Operational Vertical Sounder. *Bull. Amer. Meteor. Soc.*, **60**, 1177-1187.
- Smith, William L.; Nagle, F. W.; Hayden, C. M., and H. M. Woolf: 1981: Vertical mass and moisture structure from TIROS-N. *Bull. Amer. Meteor. Soc.*, **62**, 388- 393.
- Sundqvist, H., E. Berge, and J. E. Kristjansson, 1989: Condensation and cloud studies



- with mesoscale numerical weather prediction model. *Mon. Wea. Rev.*, **117**, 1641–1757.
- Thiébaux, J., E. Rogers, W. Wang, and B. Katz, 2003: A New High-Resolution Blended Real-Time Global Sea Surface Temperature Analysis. *Bull. Amer. Meteor. Soc.*, **84**, 645–656.
- Tiedtke, M., 1983: The sensitivity of the time-mean large-scale flow to cumulus convection in the ECMWF model. *ECMWF Workshop on Convection in Large-Scale Models*, 28 November–1 December 1983, Reading, England, pp. 297–316.
- Trenberth, K.E., and L. Smith, 2005: The Mass of the Atmosphere: A Constraint on Global Analyses. *J. Climate*, **18**, 864–875.
- Uppala, S.M., Kållberg, P.W., Simmons, A.J., Andrae, U., da Costa Bechtold, V., Fiorino, M., Gibson, J.K., Haseler, J., Hernandez, A., Kelly, G.A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R.P., Andersson, E., Arpe, K., Balmaseda, M.A., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B.J., Isaksen, L., Janssen, P.A.E.M., Jenne, R., McNally, A.P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N.A., Saunders, R.W., Simon, P., Sterl, A., Trenberth, K.E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J. 2005: The ERA-40 re-analysis. *Quart. J. R. Meteor. Soc.*, **131**, 2961–3012.
- Van den Dool, H. M., S. Saha, J. Schemm, and J. Huang, 1997: A temporal interpolation method to obtain hourly atmospheric surface pressure tides in Reanalysis 1979–1995, *J. Geophys. Res.*, **102**(D18), 22,013–22,024.
- Velicogna, I., J. Wahr, and H. Van den Dool, 2001: Can surface pressure be used to remove atmospheric contributions from GRACE data with sufficient accuracy to recover hydrological signals? *J. Geophys. Res.*, **106**(B8), 16415–16434.

- Weaver, A.T., and P. Courtier, 2001: Correlation modeling on the sphere using a generalized diffusion equation. *Quart. J. R. Meteor. Soc.*, **127**, 1815-1846.
- Wickert, J., and co-authors, 2001: Atmosphere sounding by GPS radio occultation: First results from CHAMP, *Geophys. Res.*, **105**, 7257-7273.
- Woollen, J.S., 1991: New NMC Operational OI Quality Control, Preprints, 9<sup>th</sup> AMS Conference on NWP, 1991.
- Woollen, J.S., E. Kalnay, L. Gandin, W. Collins, S. Saha, R. Kistler, M. Kanamitsu, and M. Chelliah, 1994: Quality Control in the Reanalysis System, Extended Abstracts, AMS 10<sup>th</sup> Conference on Numerical Weather Prediction, Portland, OR.
- Woolnough, S. J., J. M. Slingo, and B. J. Hoskins, 2000: The Relationship between Convection and Sea Surface Temperature on Intraseasonal Timescales. *J. Climate*, **13**, 2086-2104.
- Wu, W.-S., R.J. Purser, and D.F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.*, **130**, 2905-2916.
- Wu, X., K. S. Moorthi, K. Okamoto, and H. L. Pan, 2005: Sea ice impacts on GFS forecasts at high latitudes. *Proceedings of the 85<sup>th</sup> AMS Annual Meeting, 8<sup>th</sup> Conference on Polar Meteorology and Oceanography*, San Diego, CA.
- Xie, P., and P.A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.*, **78**, 2539 – 2558.
- Xie, P., M. Chen, A. Yatagai, T. Hayasaka, Y. Fukushima, and S. Yang, 2007: A gauge-based analysis of daily precipitation over East Asia. *J. Hydrometeor.*, **8**, 607 – 626.
- Xie, P., M. Chen, and W. Shi, 2010: CPC unified gauge analysis of global daily precipitation. To be submitted to *J. Hydrometeor.*
- Xu, K.M., and D.A. Randall, 1996: A semiempirical cloudiness parameterization for use in climate models. *J. Atmos. Sci.*, **53**, 3084-3102.
- Yu, T.-W., M. D. Iredell, and D. Keyser 1997: Global data assimilation and forecast

- experiments using SSM/I wind speed data derived from a neural network algorithm. *Wea. and Forecasting*, **12**, 859-865.
- Zhao, Q. Y., and F. H. Carr, 1997: A prognostic cloud scheme for operational NWP models. *Mon. Wea. Rev.*, **125**, 1931-1953.
- Zou, C.-Z., M. Goldberg, Z. Cheng, N. Grody, J. Sullivan, C. Cao, and D. Tarpley, 2006: Recalibration of microwave sounding unit for climate studies using simultaneous nadir overpasses. *J. Geophys. Res.* **111**, D19114: doi: 10.1029/2005JD006798, issn: 0148–0227.
- Zou, C.-Z., M. Gao, M. Goldberg, 2009: Error structure and atmospheric temperature trend in observations from the microwave sounding unit, *J. of Climate*, **22**, 1661–1681.

## LEGEND OF FIGURES

Figure 1: Diagram illustrating CFSR data dump volumes, 1978-2009, in Gb/month.

Figure 2: Performance of 500mb RADIOSONDE temperature observations. The top panel shows monthly RMS and mean fits of quality controlled observations to the first guess (blue) and the analysis (green). The fits of all observations, including those rejected by the QC, are shown in red. The bottom panel shows the 00z data counts of all observations (in red) and those which passed QC and were assimilated in green.

Figure 3: Same as Figure 2, but for 500mb RADIOSONDE wind data.

Figure 4: Same as Figure 2, but for AIRCRAFT wind data.

Figure 5: Same as Figure 2, but for ACARS wind data.

Figure 6: Same as Figure 2, but for SYNOP surface pressure data.

Figure 7: Same as Figure 2, but for METAR surface pressure data.

Figure 8: Same as Figure 2, but for MARINE temperature data.

Figure 9: Same as Figure 2, but for MARINE wind data.

Figure 10: Same as Figure 2, but for PAOB surface pressure data.

Figure 11: Same as Figure 2, but for METEOSAT satellite wind data.

Figure 12: Same as Figure 2, but for GMS satellite wind data.

Figure 13: Same as Figure 2, but for GOES satellite wind data.

Figure 14: Same as Figure 2, but for ERS ocean surface wind data.

Figure 15: Radiance instruments included in CFSR and the time period each were assimilated.

Figure 16: Performance of HIRS/2 channel 12, which peaks at ~ 500 hPa. The top panel has the uncorrected rms (o-g) in red, bias corrected (o-g) rms and mean in blue, and bias corrected (o-a) rms and mean in green, where o is obs, a is analysis and g is guess. Bottom panel as total 0Z counts in red and accepted counts in green.

Figure 17: Same as Figure 2, but for MSU channel 3, which peaks at ~ 300 hPa.

Figure 18: Same as Figure 2, but for SSU channel 1, which peaks at ~ 15 hPa.

Figure 19: Global average Tb first guess for MSU channels 1-4, with satellites denoted.

Figure 20: Same as Figure 2, but for HIRS/3 channel 5, which peaks at ~ 500 hPa.

- Figure 21: Same as Figure 2, but for AMSU-a channel 5, which peaks at ~ 500 hPa.
- Figure 22: Same as in Figure 2, but for AIRS channel 215 (channel 92 in the 281 subset) which peaks ~ 450 hPa.
- Figure 23: Comparisons of the SSU brightness temperature at channel 1 (left), 2 (middle), and 3 (right) between calculations and measurements for November 2004.
- Figure 24: On the left is the static, zonal invariant, 500 hPa stream function ( $1e^6$ ) background error valid 2007110600. On the right is the flow dependent adjusted background standard deviation.
- Figure 25: The cloud liquid water bias term for AMSUA-NOAA15 channel 3 varies with time during 2007042218 to 2007052218. The red curve initialized from zero and the blue curve is the operational values used as a reference here.
- Figure 26: The global total bias for HIRS3-NOAA17 channels 2-14 and AMSUA-NOAA15 channel 3, 5, 7, 9 varies with time during 2006110118 to 2007013018. The blue curve is the experiment values and the red curve is the operational values used as a reference.
- Figure 27: TOVS period, 1979-1998, 4xdaily averaged, globally averaged, total bias correction for MSU channels 1-4 (left) and SSU channels 1-3 (right).
- Figure 28: The yearly total of tropical storm reports stacked by the 8 geographical basins from top to bottom, Western Pacific (W Pac), Southern Indian (S Ind), Southern Pacific (S Pac) , North Atlantic-Caribbean (Atl), Eastern Pacific (E Pac), Central Pacific (Cen Pac) , Bay of Bengal and Arabian Sea.
- Figure 29: Time series of the percentage of detected tropical storms plotted globally and for selected Northern Hemisphere basins: the Atlantic-Caribbean (ATL), Western Pacific (W Pac) and Eastern Pacific (E Pac) Oceans.
- Figure 30: The vertical structure of model levels as a meridional cross section at 90E. The left panels are for R1 (28 sigma layers) and right panels for CFSR (64 sigma-pressure hybrid layers). The top panels are plotted as a linear function of pressure to emphasize resolution in the troposphere and bottom panels are plotted in log(pressure) to emphasize stratosphere.
- Figure 31: Same as in Figure 30, but as a zonal cross section across Rockies at 40N.
- Figure 32: The global number of temperature observations assimilated per month by the

ocean component of the CFSR as a function of depth for the years 1980 through 2008. The contour interval is 250 observations.

Figure 33: The global distribution of all temperature profiles assimilated by the ocean component of the CFSR for the year 1985

Figure 34: The same as Figure 33, but for the year 2008.

Figure 35: Sea ice concentration for September of 1987 (top) and 2007 (bottom) for the Arctic.

Figure 36: Monthly mean Sea ice extent ( $10^6 \text{ km}^2$ ) for the Arctic (top) and the Antarctic (bottom).

Figure 37: CFS grid architecture.

Figure 38: 2-meter volumetric soil moisture climatology of CFSR for May averaged over 1980-2008.

Figure 39: Same as in Figure 38, except for November.

Figure 40: Schematic of the execution of one day of the CFS Reanalysis.

Figure 41: In the upper diagram are the daily time series of 0000 GMT 120 hour forecasts for the T126L64 CFSR (black), the T62L64 CFSR-Lite (red) , and operational (OPR) T382L64 operational GFS for 2006. The lower diagram is the time series of the difference CFSR minus GFS.

Figure 42: Yearly averaged Southern Hemisphere (left) and Northern Hemisphere (right) 00 GMT 120 hour forecast anomaly correlations for CFSR (black triangles), GFS (red circles), CFSR-Lite (green squares), CFS R2 (purple diamonds), CDAS R1 (blue stars).

Figure 43: Global average of monthly-mean precipitation (a), evaporation (b) and evaporation minus precipitation (c). Averages over ocean, land, and ocean plus land are plotted with red, blue, and black curves. Unit is mm/day.

Figure 44: Monthly mean hourly surface pressure (with the daily mean subtracted out) for March 1998 in a collage of 24 global maps. Each map is on the T382 Gaussian grid. In red (blue) areas pressure is higher (lower) than the daily mean. Units are hPa. Time starts in the upper left (00Z), then proceeds down to 3Z, then continues at the top of the next column of global maps etc.

Figure 45: The fit of 6-hour forecasts of instantaneous surface pressure against irregularly distributed observations also called the background error. Shown are annually compiled fits-to-obs data 1979-2008 in units of mbar. In blue SH ocean, in red NH land.

Figure 46: Global mean temperature anomalies from 1000 to 1 hPa from January 1979 through May 2009.

Figure 47: Equatorial (5°S to 5°N) zonal mean U wind component time series from 1979 to 2009.

Figure 48: Zonal mean total ozone differences between the CFSR and the Cohesive SBUV(/2) Total Ozone Dataset. Units are in Dobson units. Also shown are The time spans of each satellite's SBUV(/2) data used in the Cohesive SBUV(/2) Total Ozone Dataset.

Figure 48: Zonal mean total ozone anomalies as a function of latitude vs time.

Figure 49: Temporal lag correlation coefficient between precipitation and SST in the tropical Western Pacific (averaged over 10S-10N, 130E-150E) in R1 (red), R2 (brown), CFSR (green) and observation (black). GPCP daily precipitation and Reynolds  $\frac{1}{4}^\circ$  daily SST are used as observational data. Negative (positive) lag days on the x-axis indicate the SST leads (lags) the precipitation. Data for the boreal winter (Nov-Apr) over the period 1979-2008 are band-pass filtered for 20-100 days after removing the climatological mean.

Figure 50: The subsurface temperature mean for an equatorial cross-section for CFSR (top) and observations from the World Ocean Atlas (Conkright et al., 1999).

Figure 51: The subsurface zonal velocity for an equatorial cross-section for CFSR for the years: 1979-2008.

Figure 52: Vertical profiles of the subsurface temperature (top panels) and zonal velocity (bottom panels) for CFSR (red line) and TAO observations (black line) for four locations in the equatorial Pacific Ocean.

Figure 53: The vertically averaged temperature (top 300 m) for CFSR for 1979-2008, and for observations from the World Ocean Atlas (Conkright, et al., 1999).

Figure 54: Zonal surface velocities for CFSR ( top left) and for drifters from the Surface

Velocity Program of TOGA (top right). Meridional velocities are shown in the bottom panel.

Figure 55: The first two EOFs of the SSH variability for the CFSR (left) and for TOPEX satellite altimeter data (right) for the period: 1993-2008. The time series amplitude factors are plotted in the bottom panel.



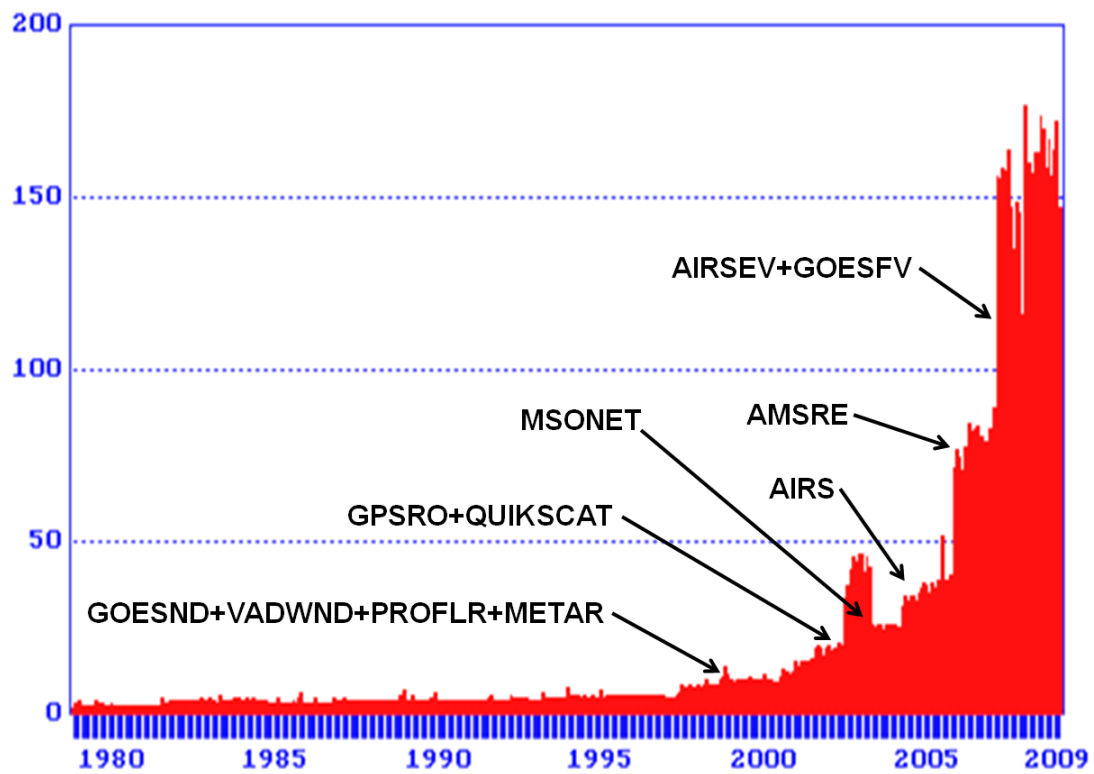


Figure 1: Diagram illustrating CFSR data dump volumes, 1978-2009, in Gb/month.

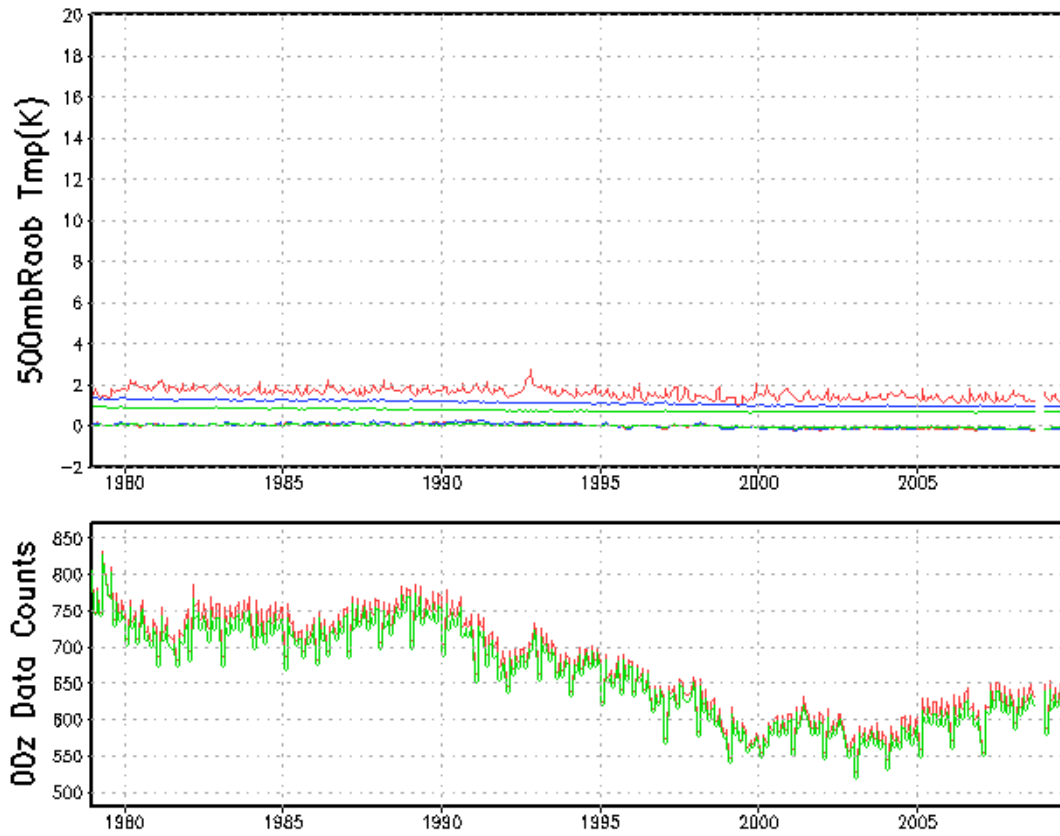


Figure 2: Performance of 500mb RADIOSONDE temperature observations. The top panel shows monthly RMS and mean fits of quality controlled observations to the first guess (blue) and the analysis (green). The fits of all observations, including those rejected by the QC, are shown in red. The bottom panel shows the 00z data counts of all observations (in red) and those which passed QC and were assimilated in green.

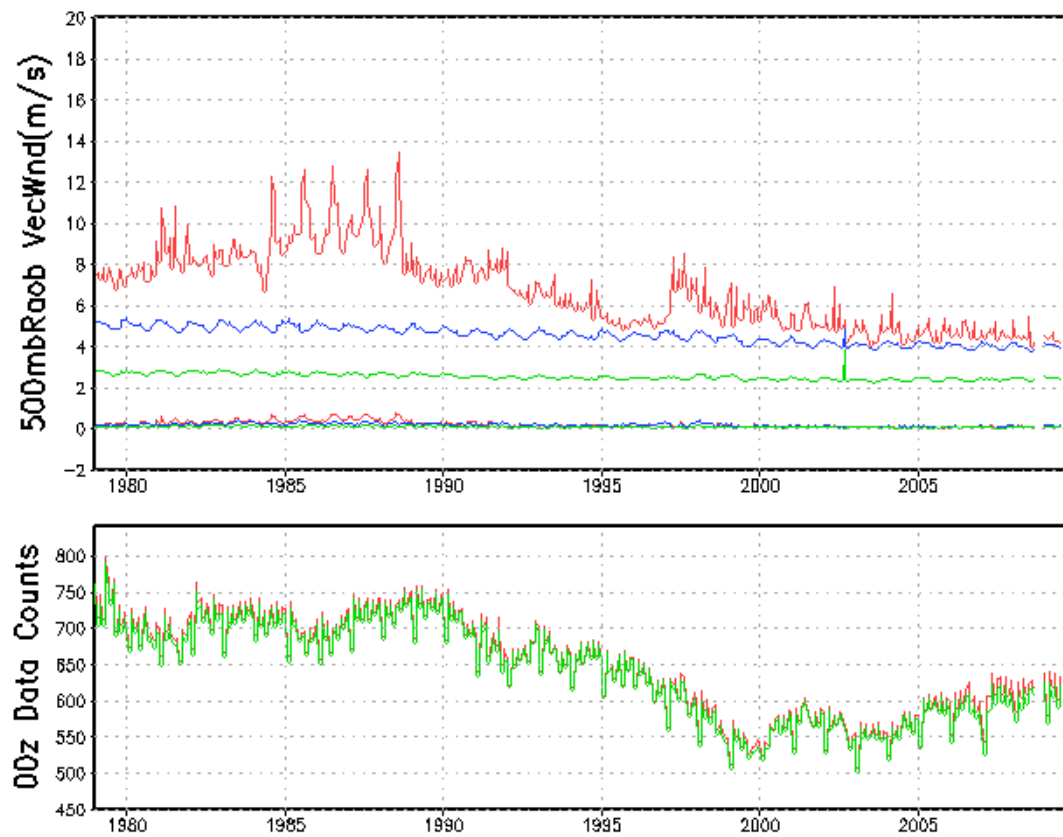


Figure 3: Same as in Figure 2, but for 500mb RADIOSONDE wind data.

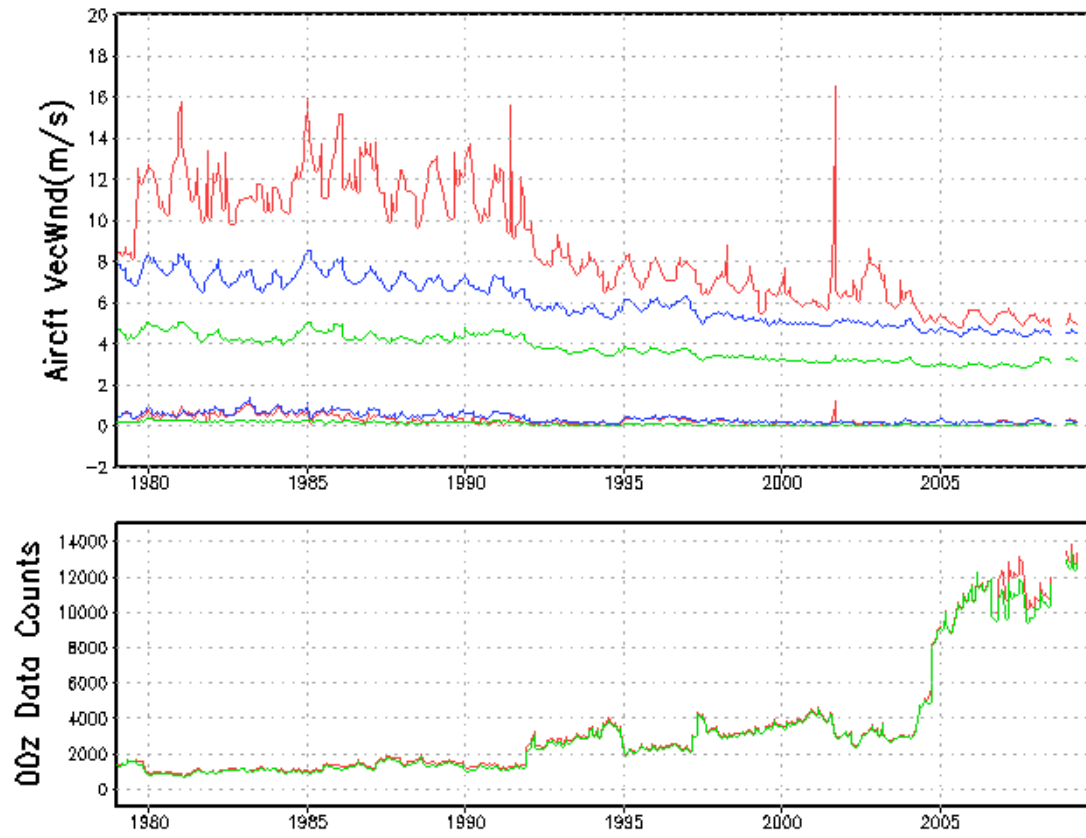


Figure 4: Same as in Figure 2, but for AIRCRAFT wind data.

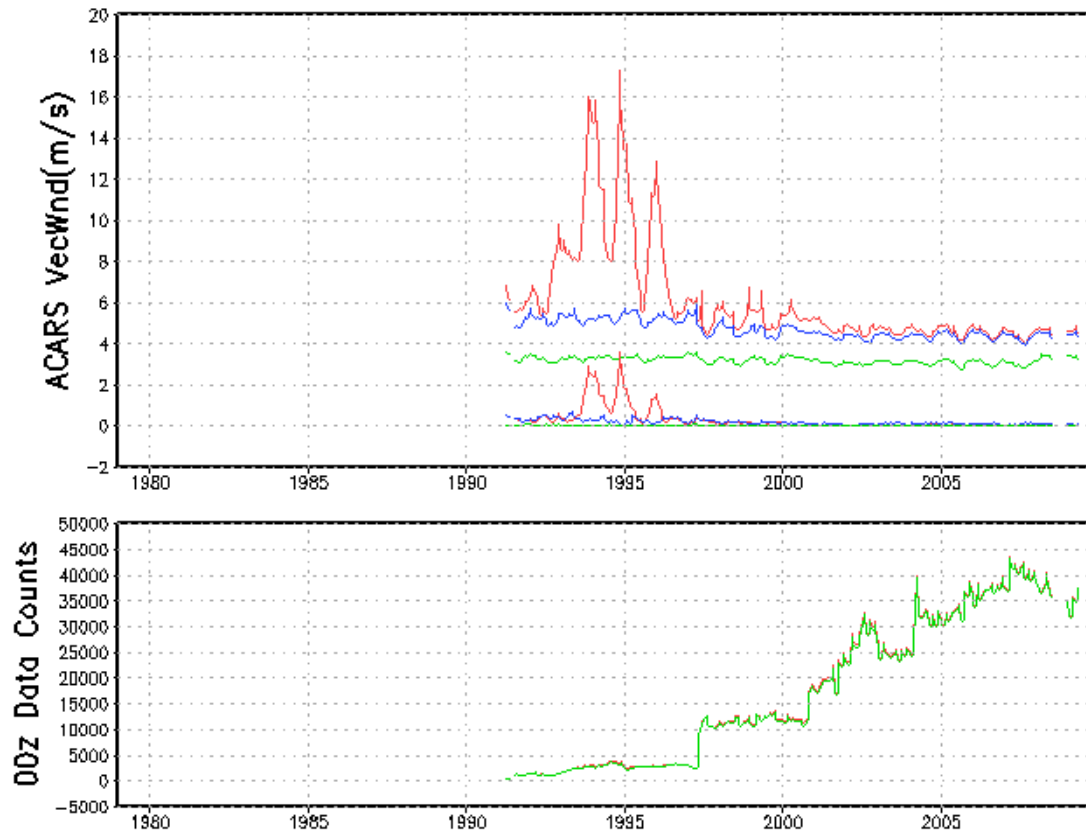


Figure 5: Same as in Figure 2, but for ACARS wind data.

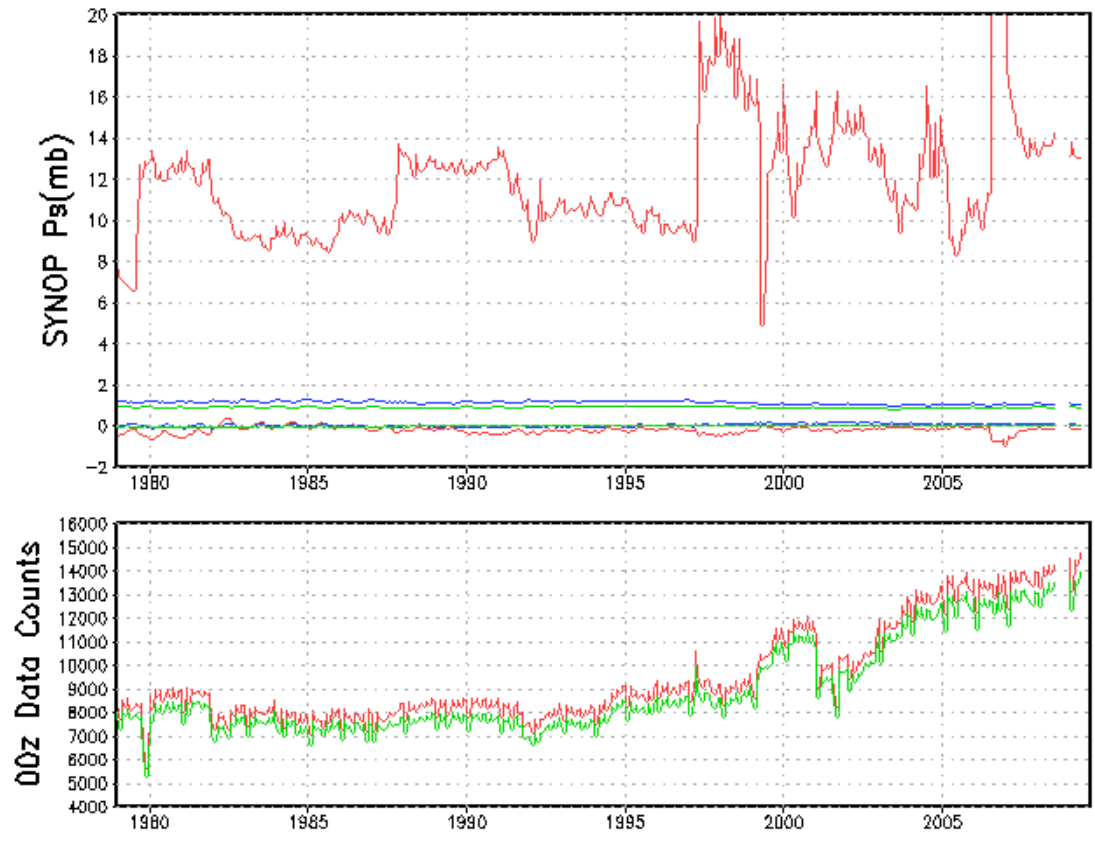


Figure 6: Same as in Figure 2, but for SYNOP surface pressure data.

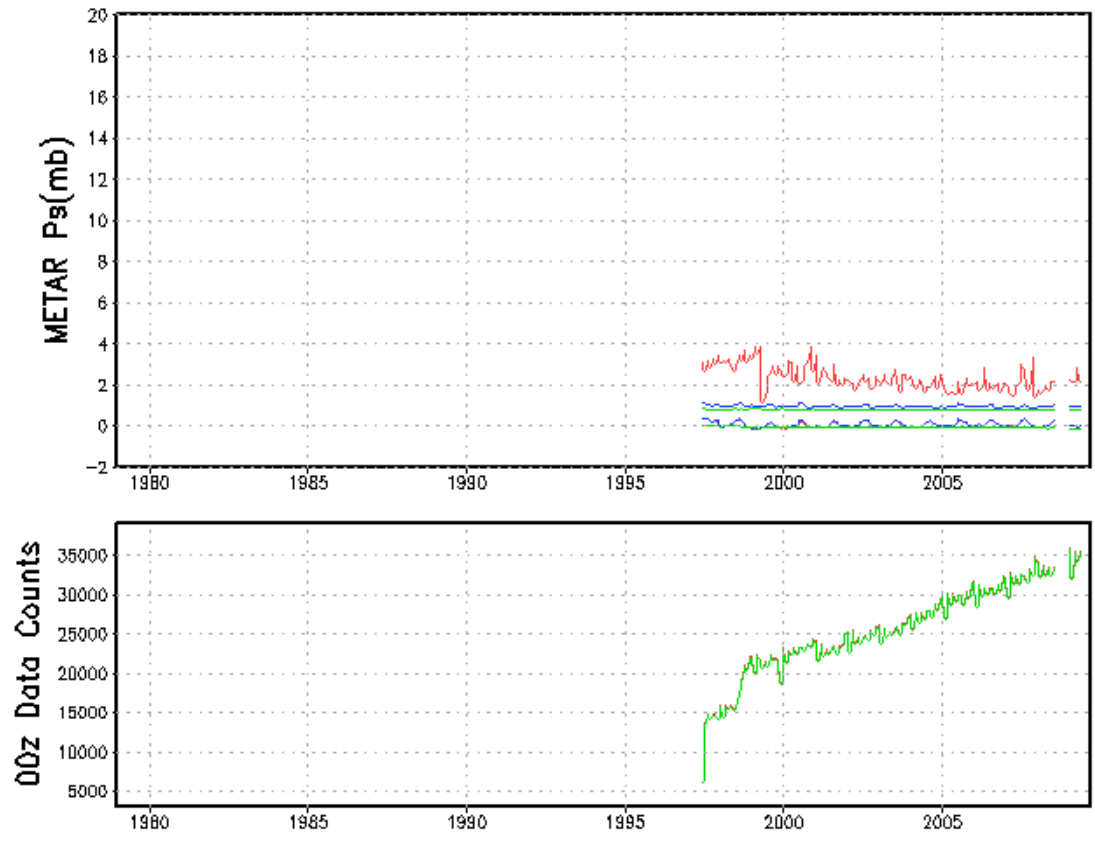


Figure 7: Same as in Figure 2, but for METAR surface pressure data

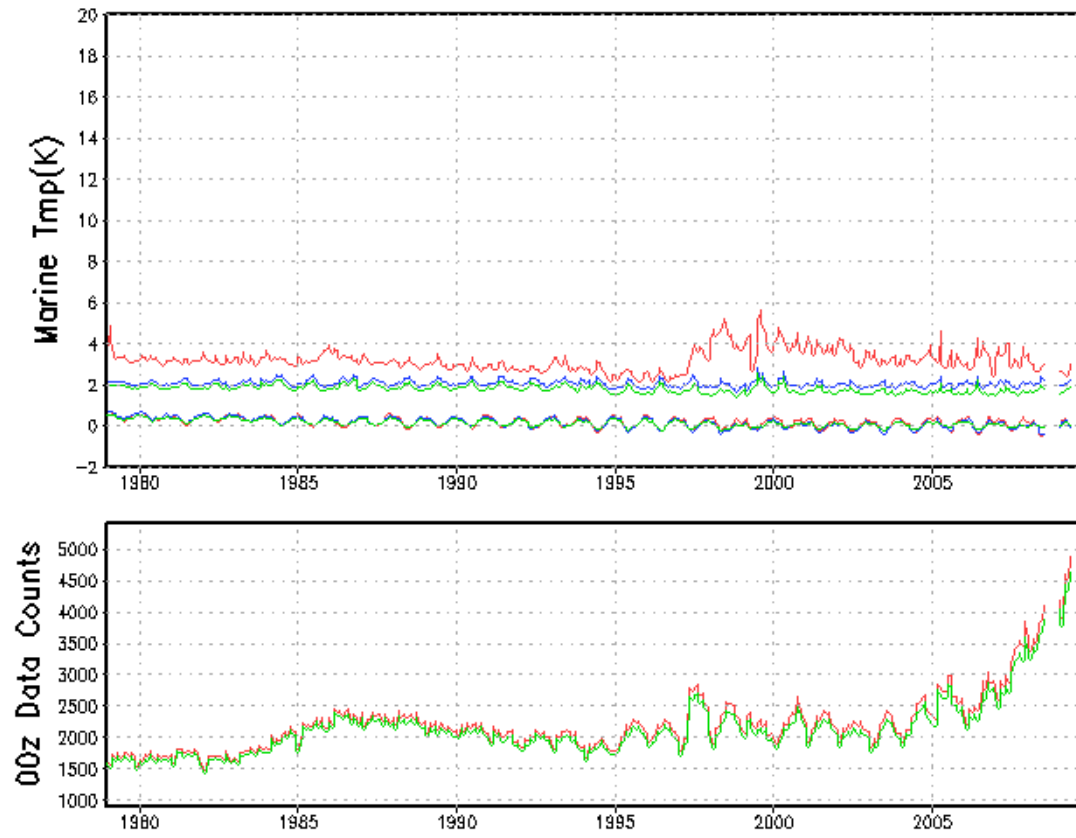


Figure 8: Same as in Figure 2, but for MARINE temperature data.



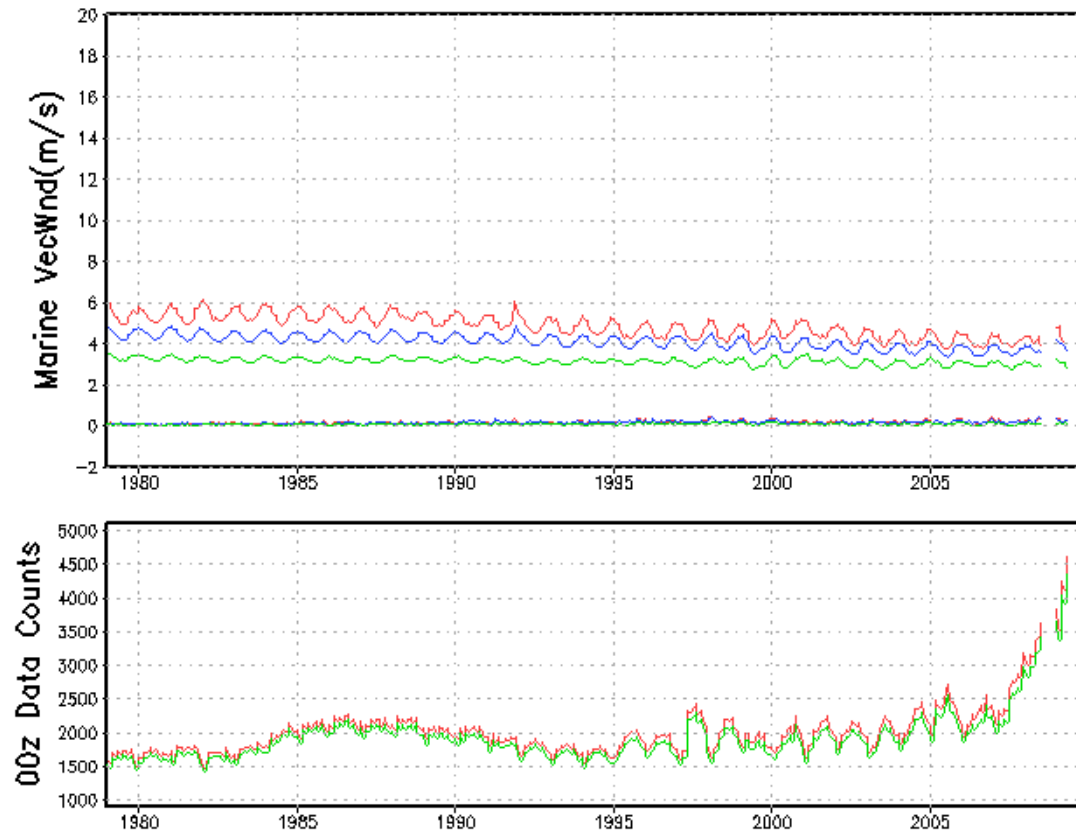


Figure 9: Same as in Figure 2, but for MARINE wind data.

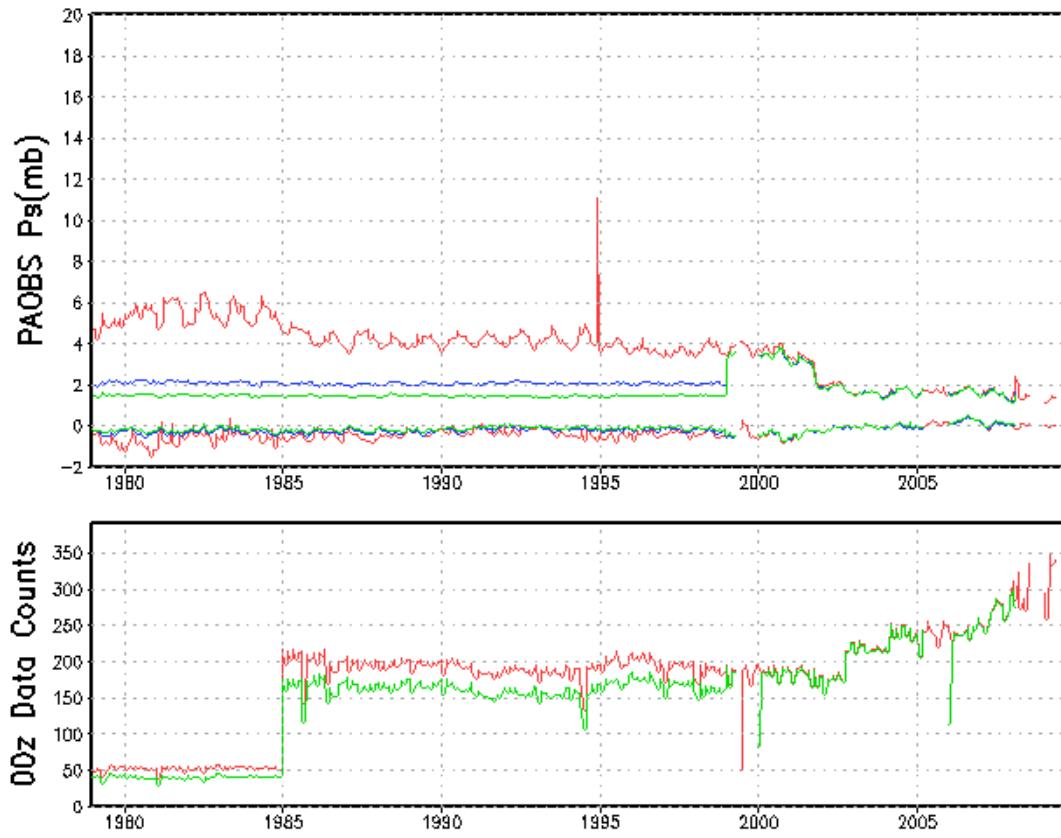


Figure 10: Same as in Figure 2, but for PAOB surface pressure data.

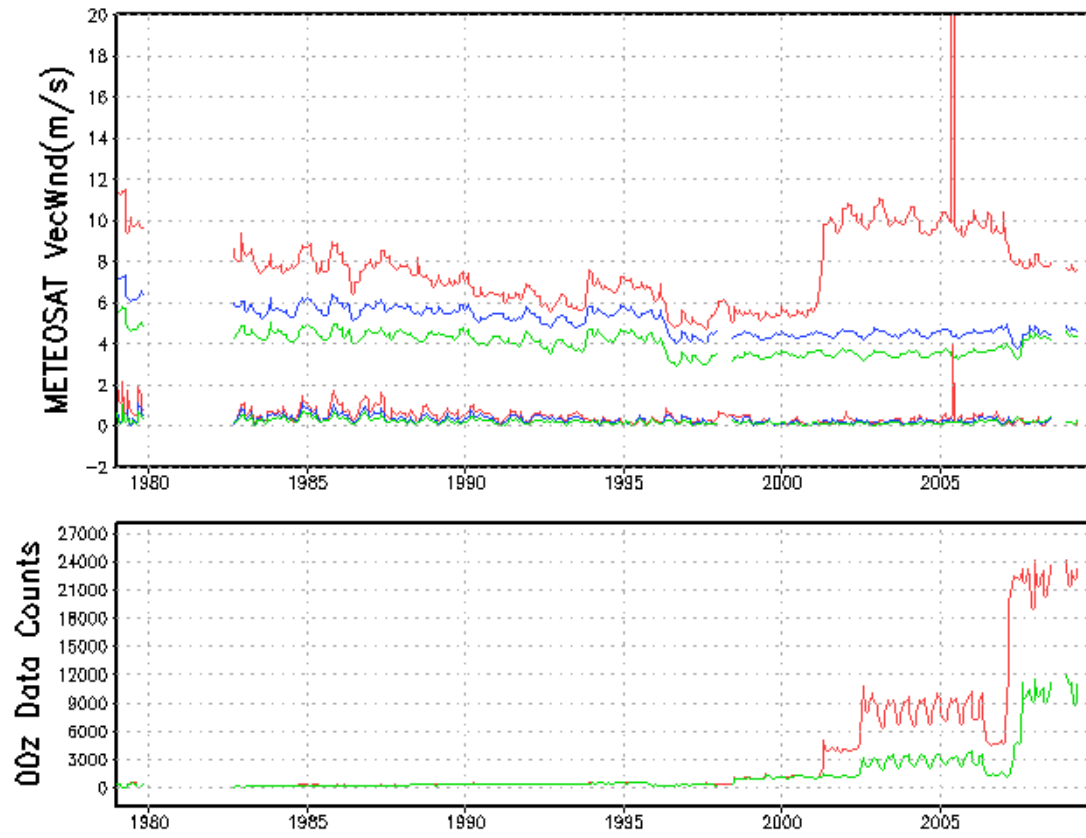


Figure 11: Same as in Figure 2, but for METEOSAT satellite wind data.

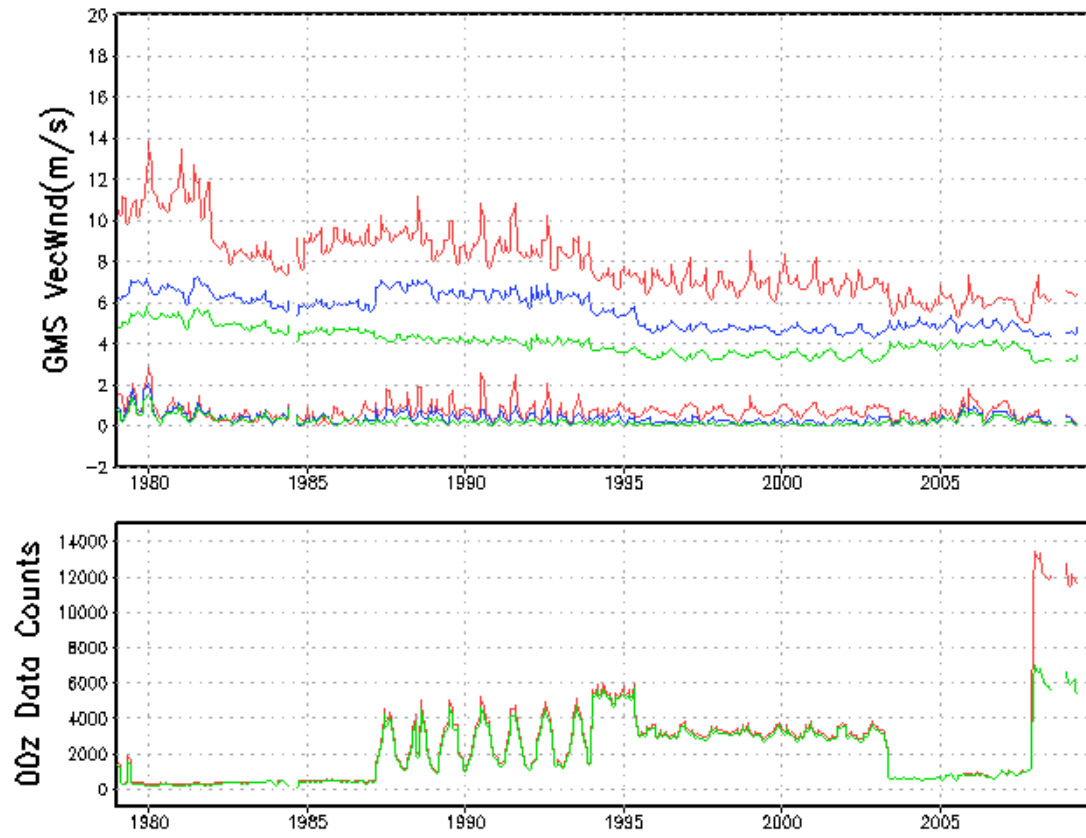


Figure 12: Same as in Figure 2, but for GMS satellite wind data.

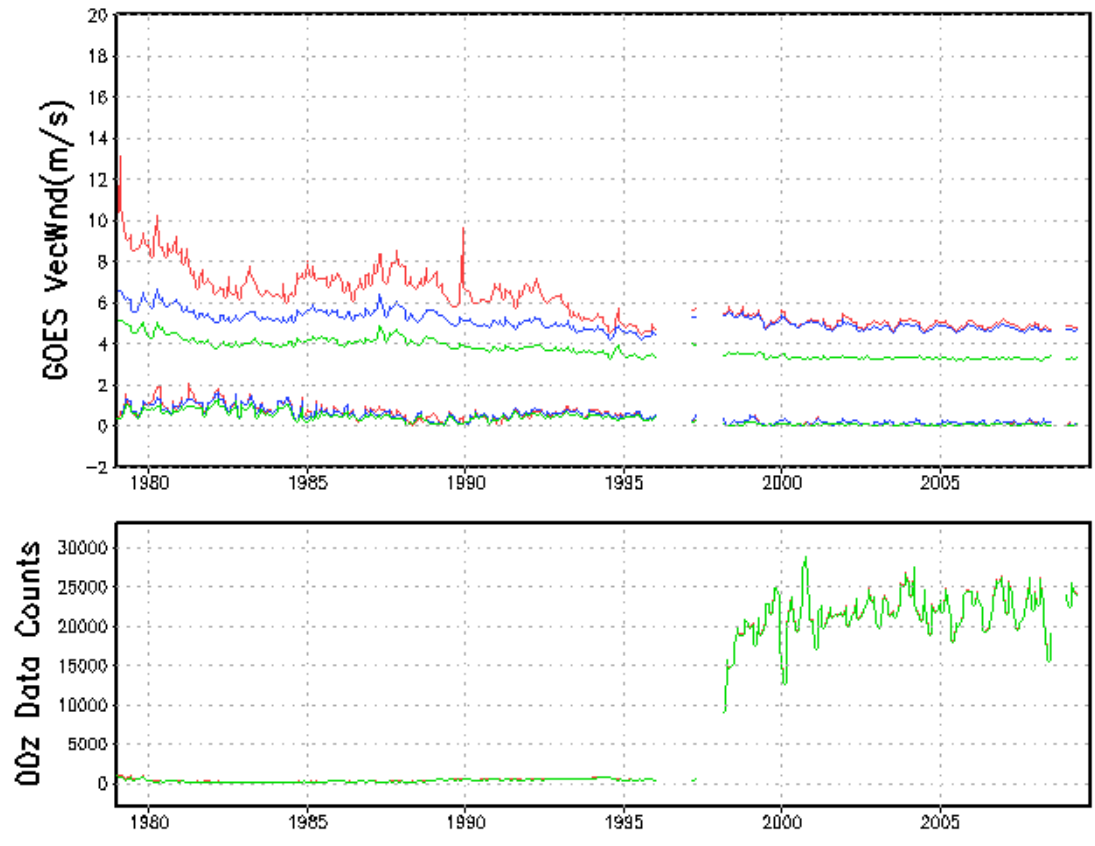


Figure 13: Same as in Figure 2, but for GOES satellite wind data.

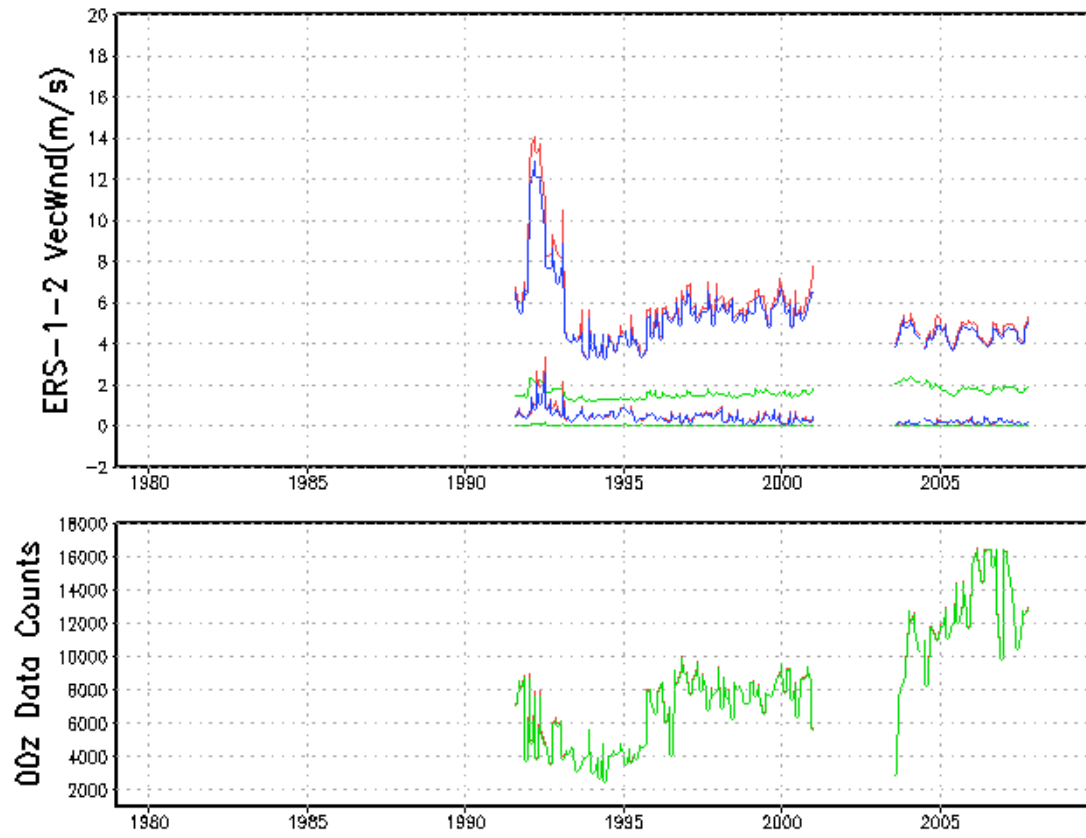


Figure 14: Same as in Figure 2, but for ERS ocean surface wind data.

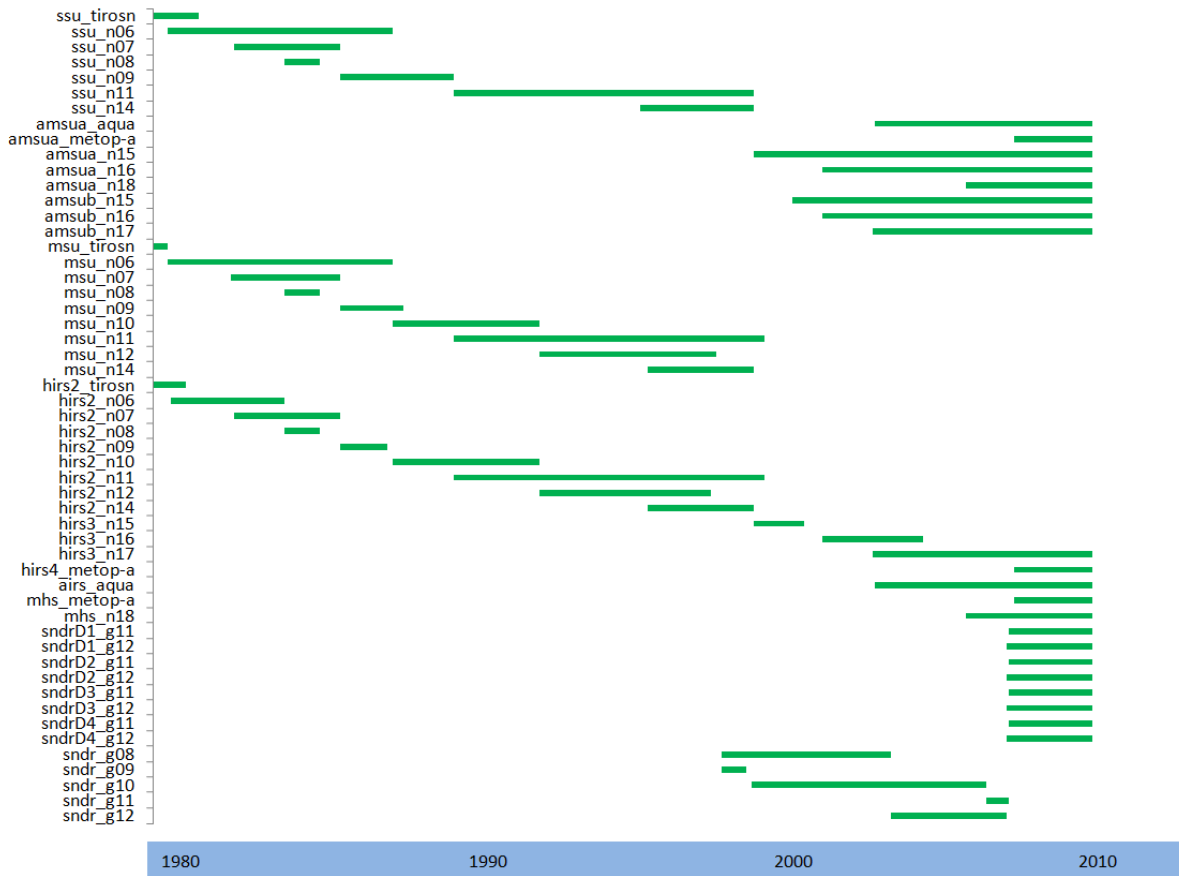


Figure 15: Radiance instruments on satellites included in CFSR and the time period each were assimilated.

Satellite references: Most of the satellite information searches start from the NCDC CLASS website: <http://www.class.ncdc.noaa.gov/saa/products/welcome>

More information on specific platforms:

TOVS and ATOVS: <http://www.ncdc.noaa.gov/oa/pod-guide/ncdc/docs/intro.htm>

GOES: <http://www.oso.noaa.gov/goes/>

DMSP: <http://heasarc.gsfc.nasa.gov/docs/heasarc/missions/dmsp.html>

AQUA/AIRS: [http://www.nasa.gov/mission\\_pages/aqua/](http://www.nasa.gov/mission_pages/aqua/)

A good overview of NOAA satellites: Davis (2007).

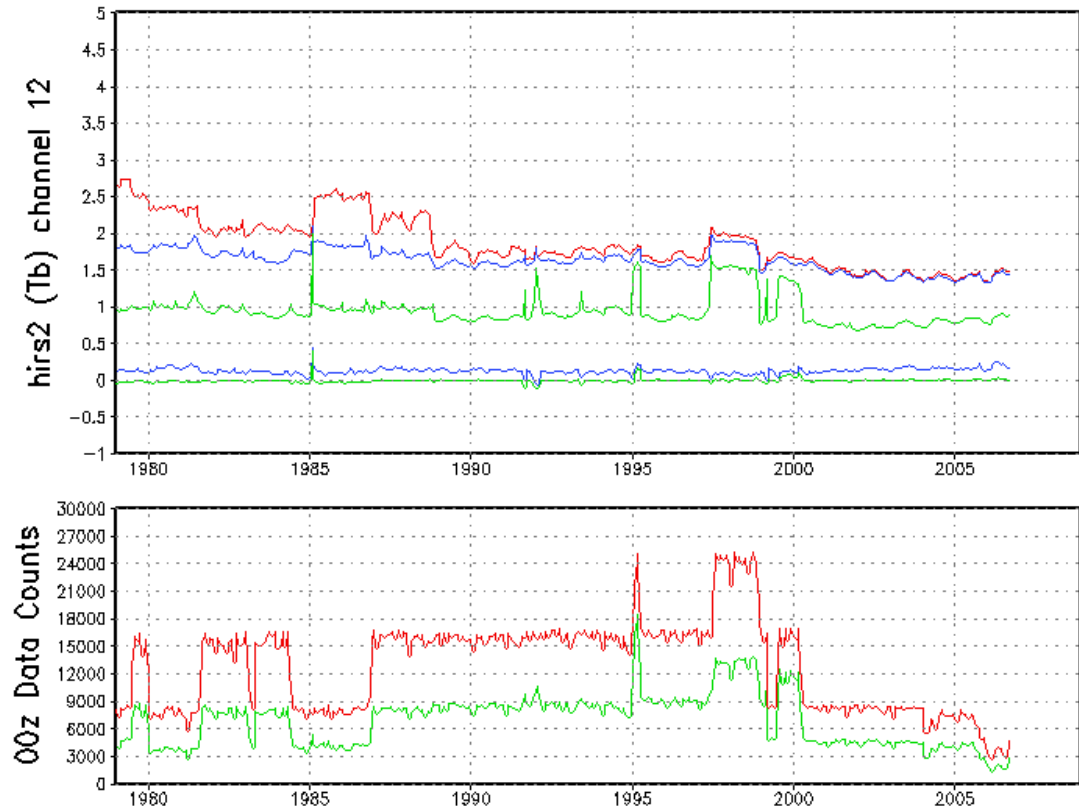


Figure 16: Performance of HIRS/2 channel 12, which peaks at  $\sim 500$  hPa. The top panel has the uncorrected rms (o-g) in red, bias corrected (o-g) rms and mean in blue, and bias corrected (o-a) rms and mean in green, where o is obs, a is analysis and g is guess. Bottom panel as total 0Z counts in red and accepted counts in green.



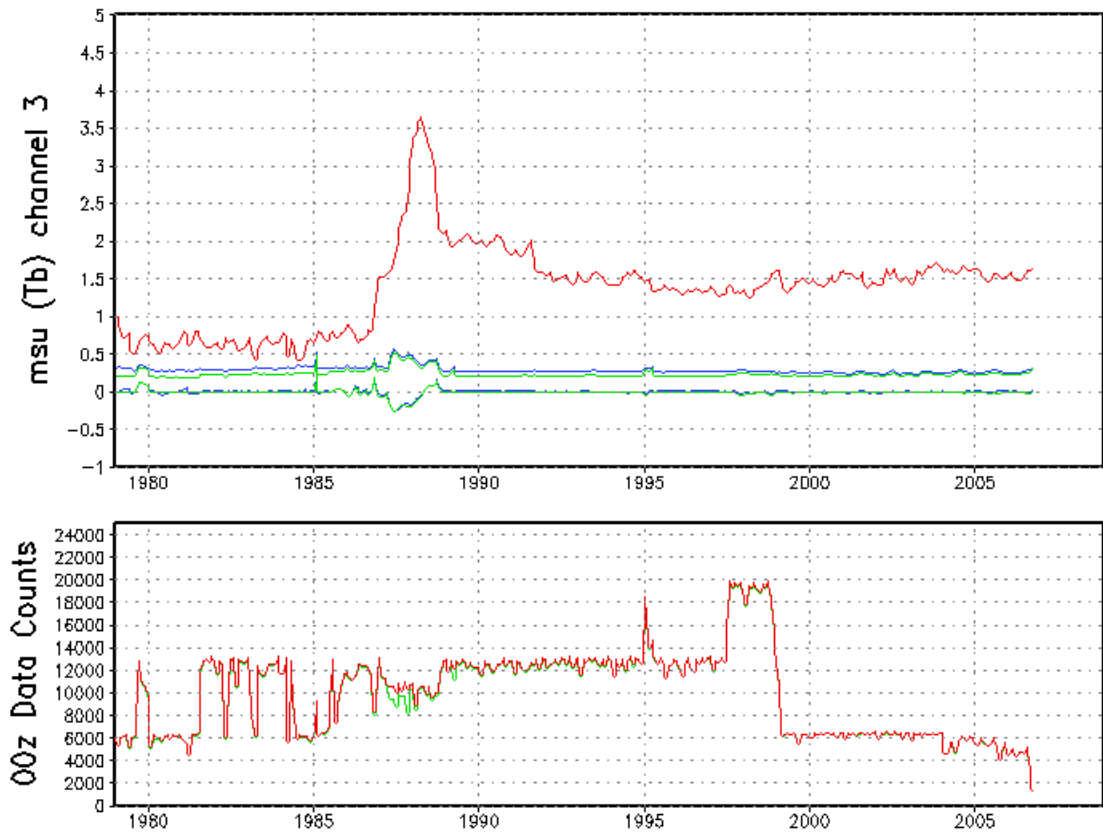


Figure 17: Same as in Figure 16, but for MSU channel 3, which peaks at ~ 300 hPa

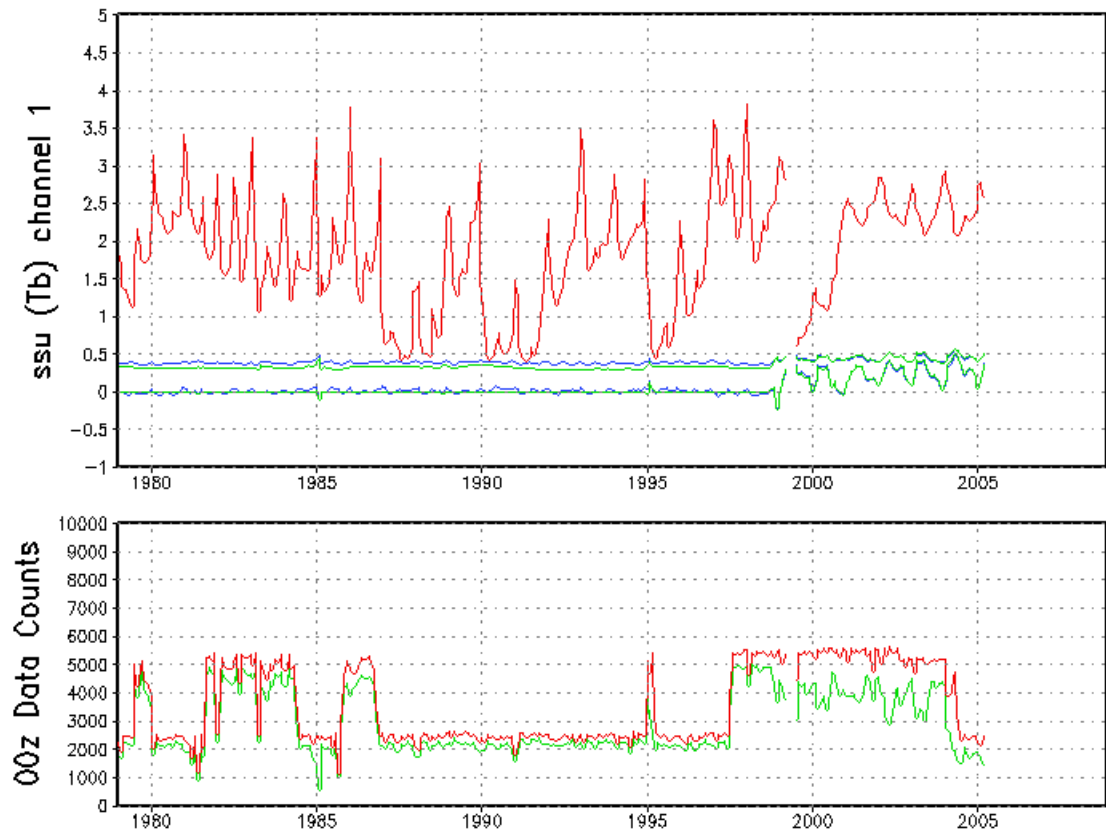


Figure 18: Same as in Figure 16, but for SSU channel 1, which peaks at  $\sim 15$  hPa.

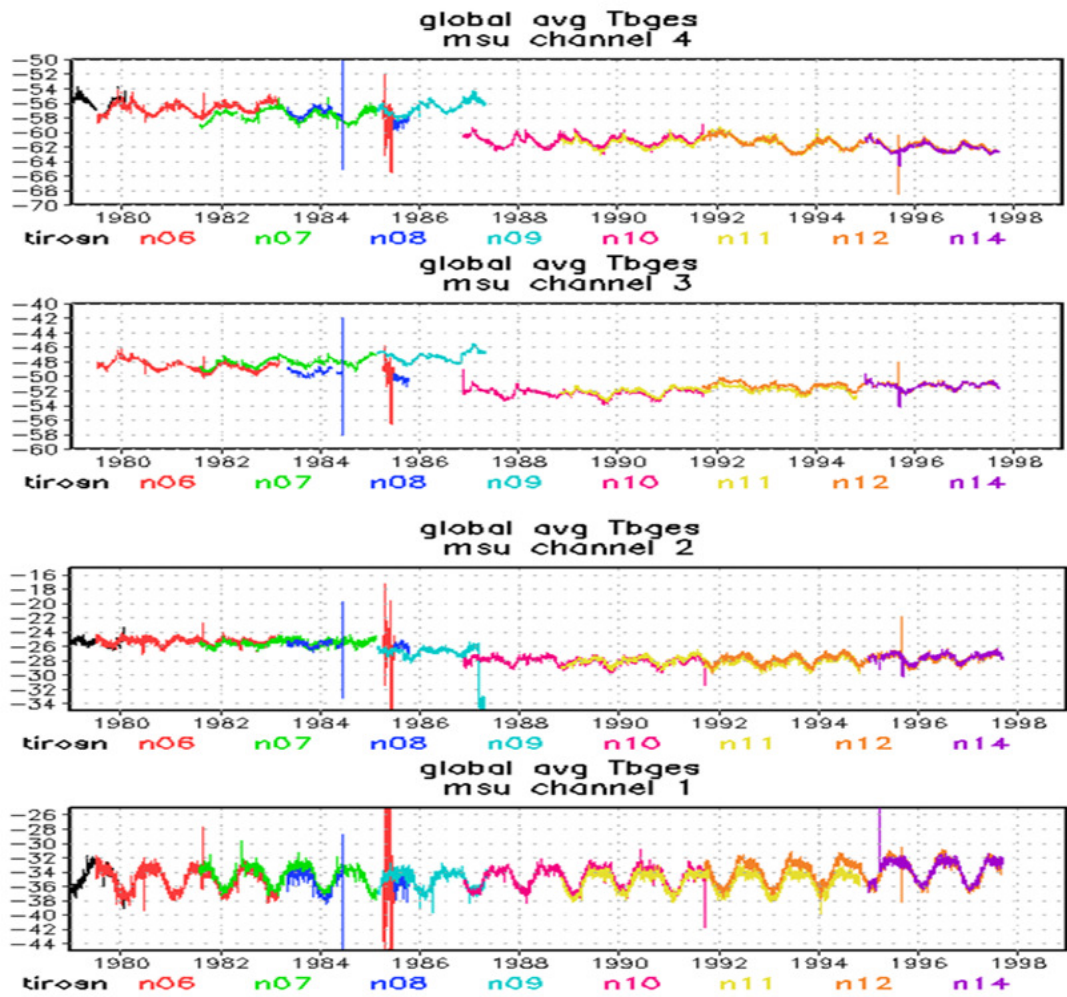


Figure 19: Global average Tb first guess for MSU channels 1-4, with satellites denoted

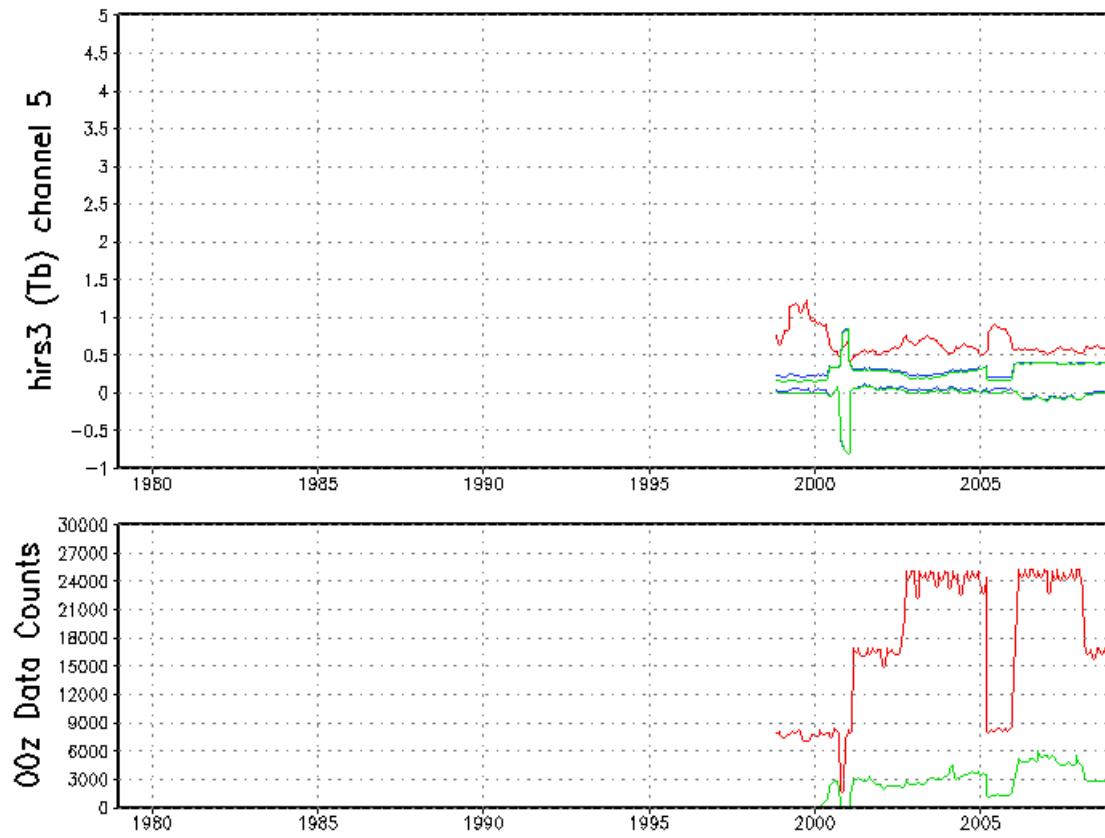


Figure 20: Same as in Figure 16, but for HIRS/3 channel 5, which peaks at  $\sim 500$  hPa.

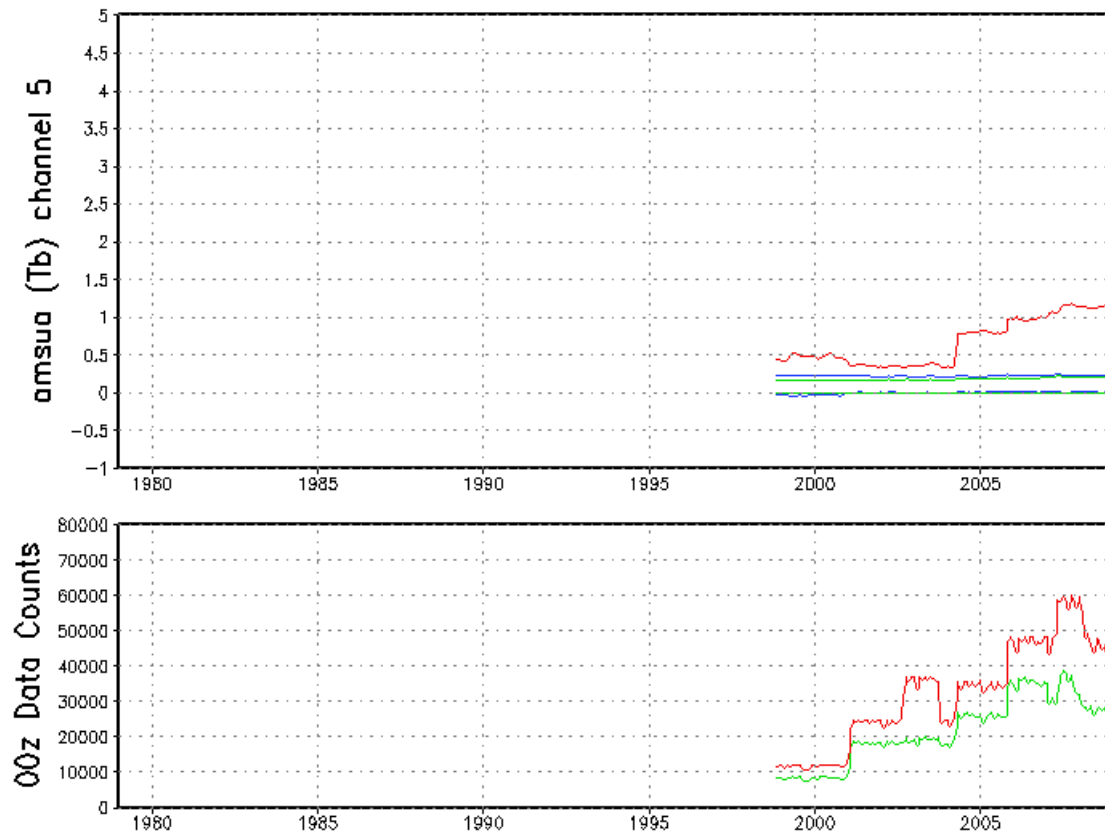


Figure 21: Same as in Figure 16, but for AMSU-a channel 5, which peaks at  $\sim 500$  hPa.

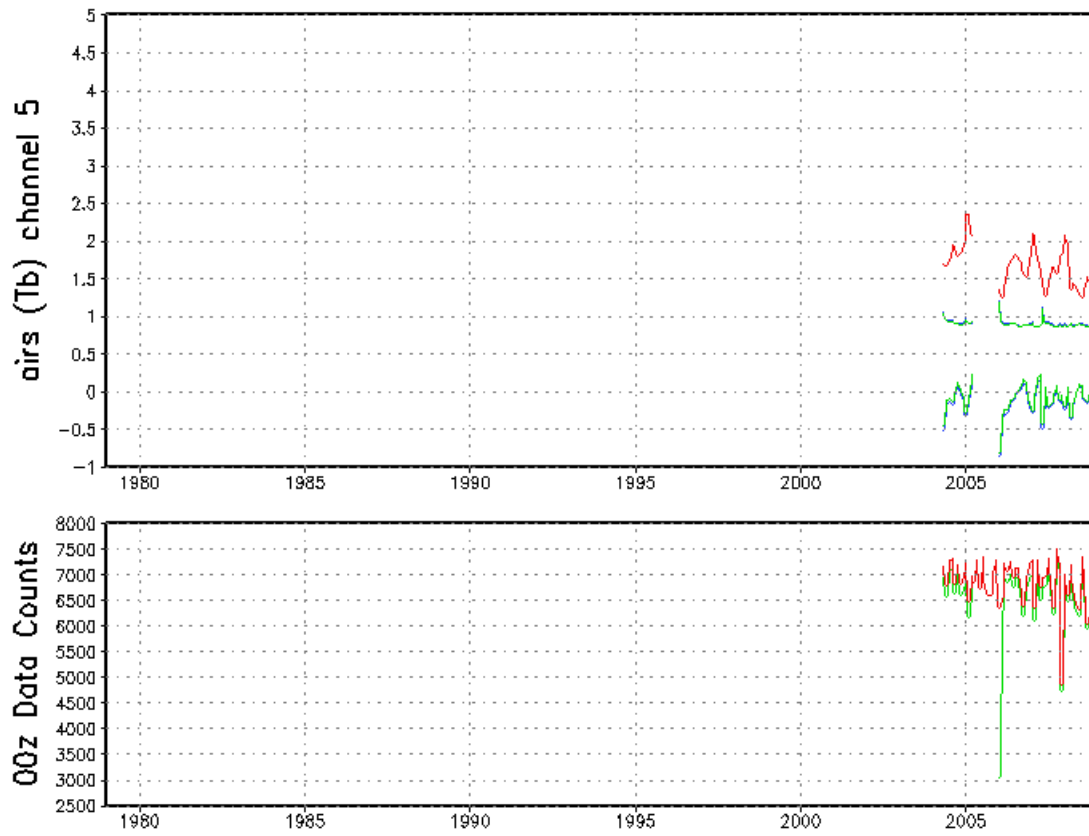


Figure 22: Same as in Figure 16, but for AIRS channel 215 (channel 92 in the 281 channel subset), which peaks at ~450hPa.

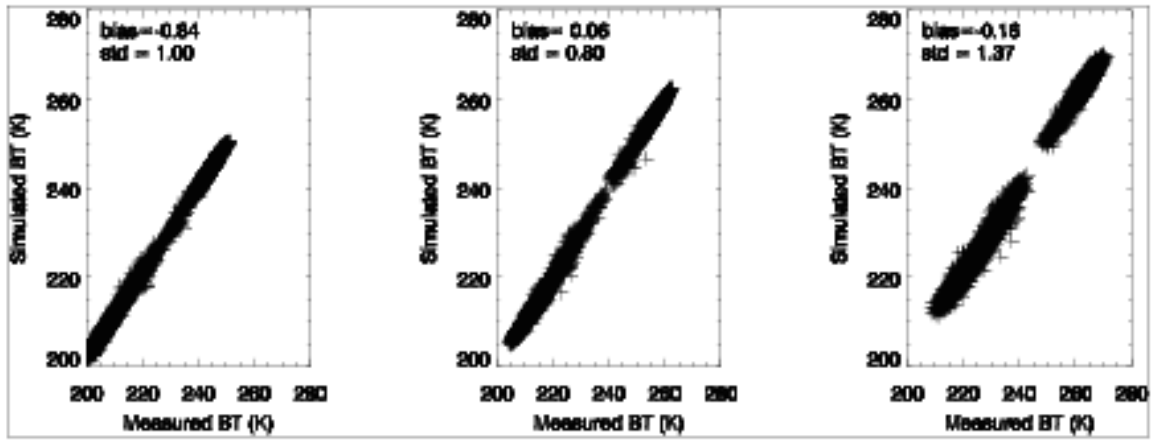


Figure 23: Comparisons of the SSU brightness temperature at channel 1 (left), 2 (middle), and 3 (right) between calculations and measurements for November 2004.

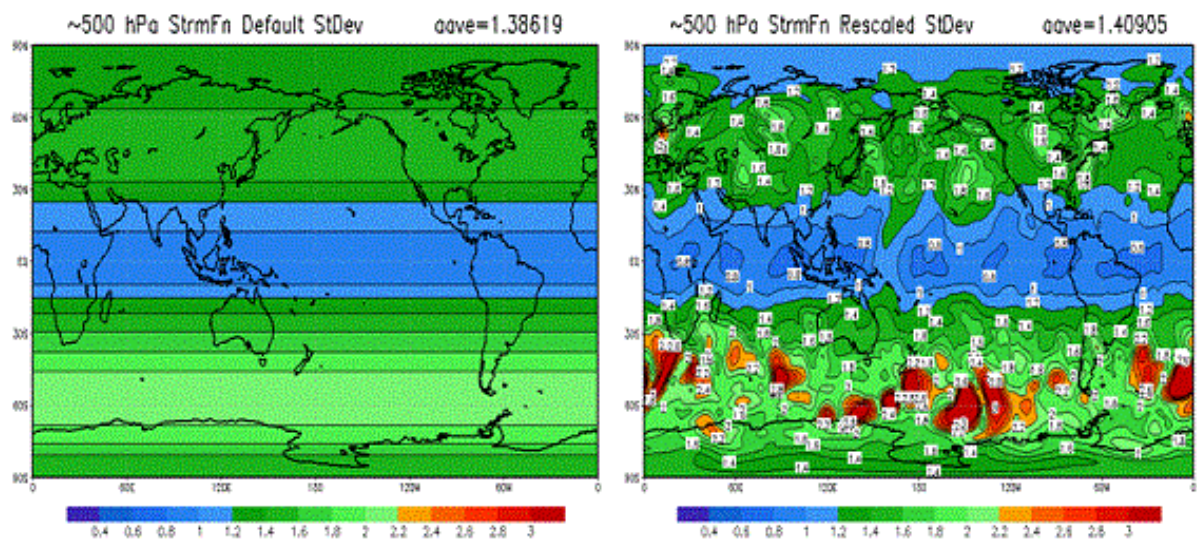


Figure 24: On the left is the static, zonal invariant, 500 hPa stream function ( $1e6$ ) background error valid 2007110600. On the right is the flow dependent adjusted background standard deviation.



platform: amsua\_n15  
region : global  
variable: cloud liquid water correction (K)  
valid : 18Z22APR2007 to 18Z22MAY2007

channel 3  
 $\chi$  0.7350  
 $f$  57.29 GHz  
 $\lambda$  5232.86  $\mu\text{m}$

file 1:  
avg: 0.8596  
sdv: 1.8824

file 2:  
avg: 0.0472  
sdv: 0.1047

file 3:  
avg: 0.6963  
sdv: 1.5138

file 4:  
avg: 0.6820  
sdv: 1.4980

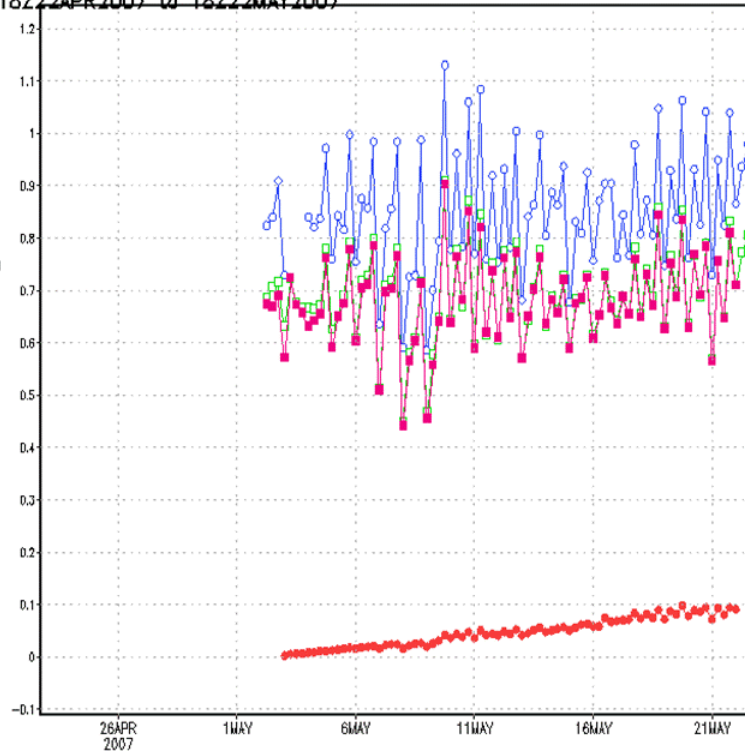
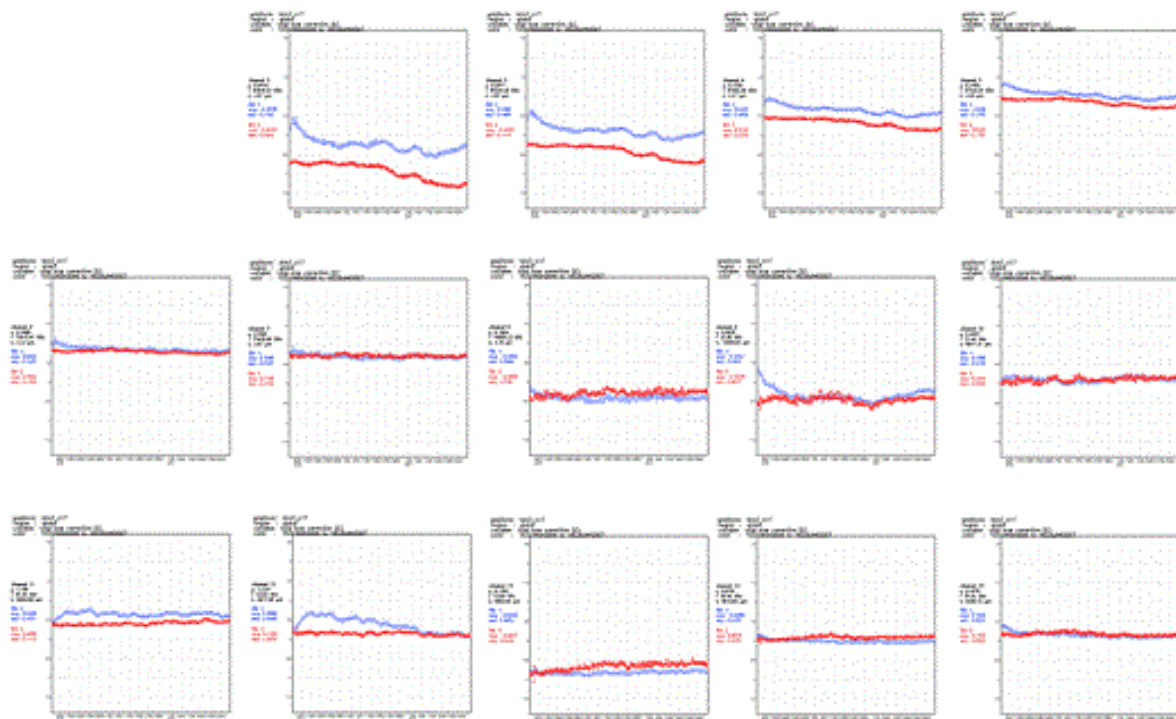
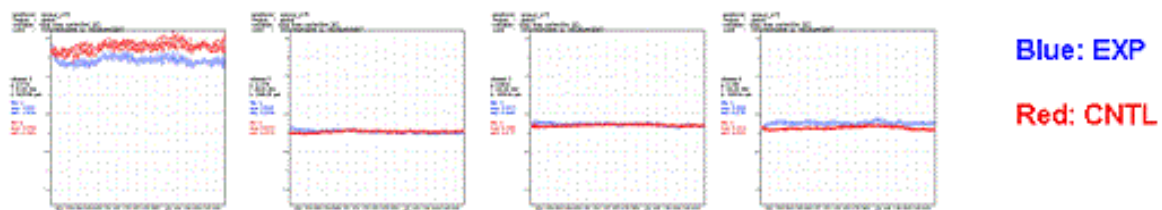


Figure 25: The cloud liquid water bias term for AMSUA-NOAA15 channel 3 varies with time during 2007042218 to 2007052218. The red curve initialized from zero and the blue curve is the operational values used as a reference here.

## Global Total Bias 2006110118 - 2007013018



## HIRS3-NOAA 17 channels 2-14



Blue: EXP

Red: CNTL

Figure 26: The global total bias for HIRS3-NOAA17 channels 2-14 and AMSUA-NOAA15 channel 3, 5, 7, 9 varies with time during 2006110118 to 2007013018. The blue curve is the experiment values and the red curve is the operational values used as a reference.

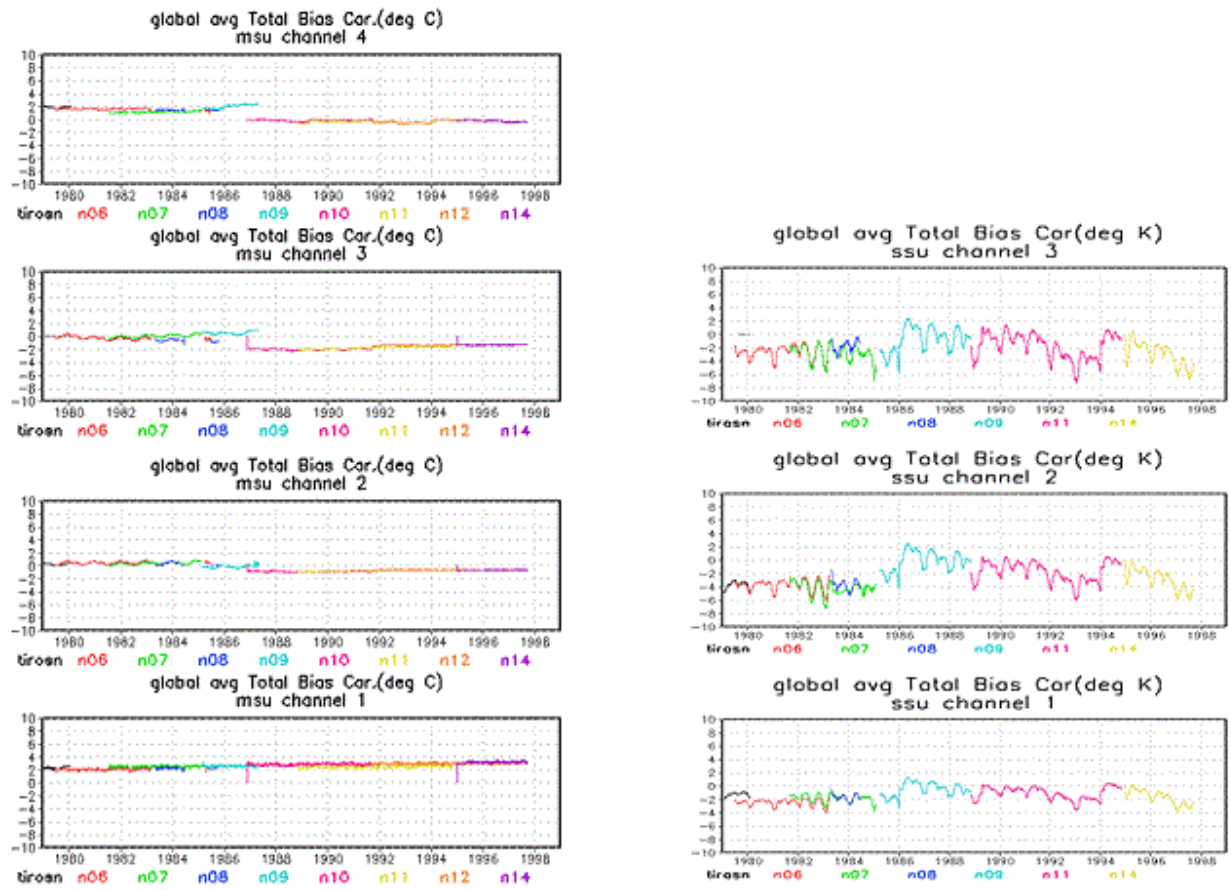


Figure 27: TOVS period, 1979-1998, 4xdaily averaged, globally averaged, total bias correction for MSU channels 1-4 (left) and SSU channels 1-3 (right).

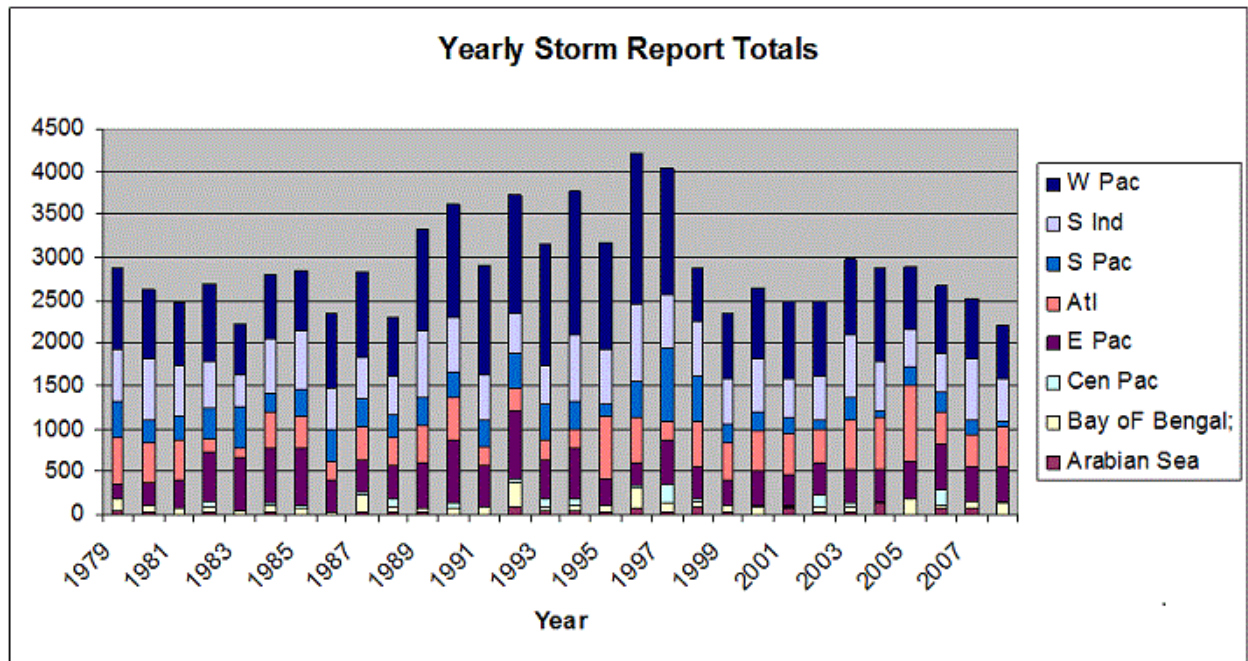


Figure 28: The yearly total of tropical storm reports stacked by the 8 geographical basins

from top to bottom, Western Pacific (W Pac), Southern Indian (S Ind), Southern Pacific (S Pac) , North Atlantic-Caribbean (Atl), Eastern Pacific (E Pac), Central Pacific (Cen Pac) , Bay of Bengal and Arabian Sea.

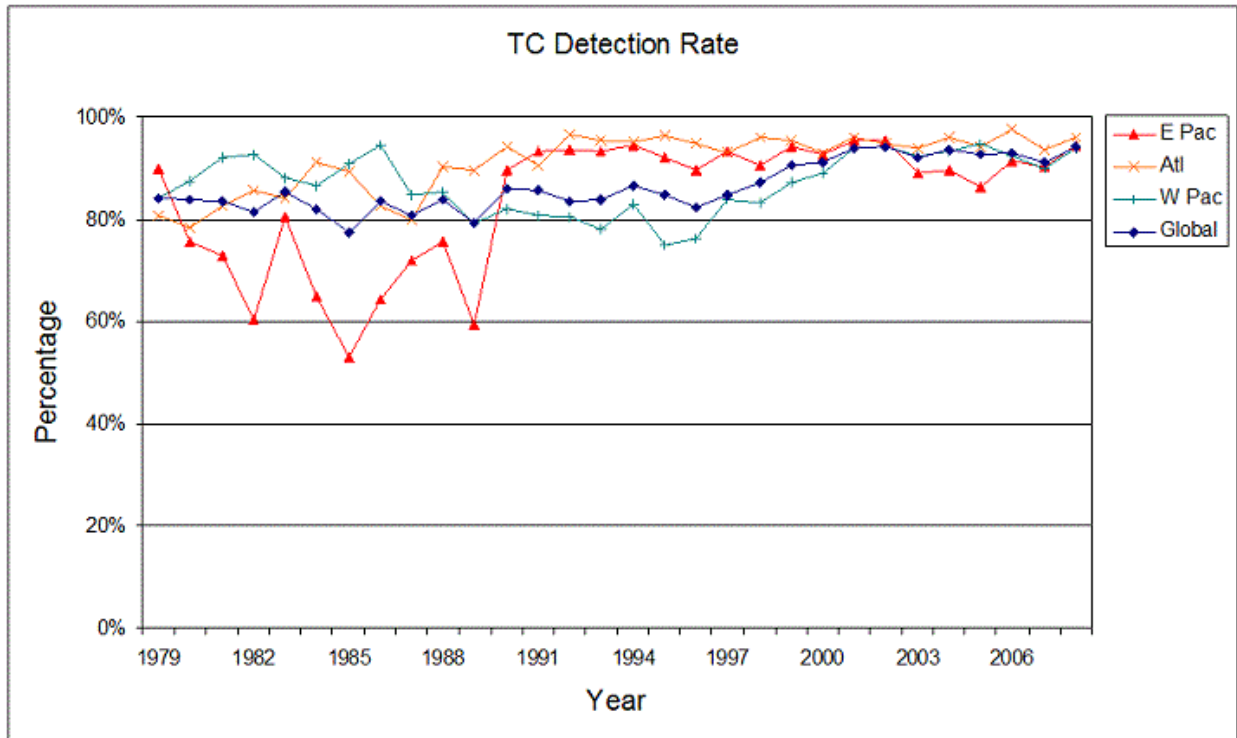


Figure 29: Time series of the percentage of detected tropical storms plotted globally and for selected Northern Hemisphere basins: the Atlantic-Caribbean (ATL), Western Pacific (W Pac) and Eastern Pacific (E Pac) Oceans.

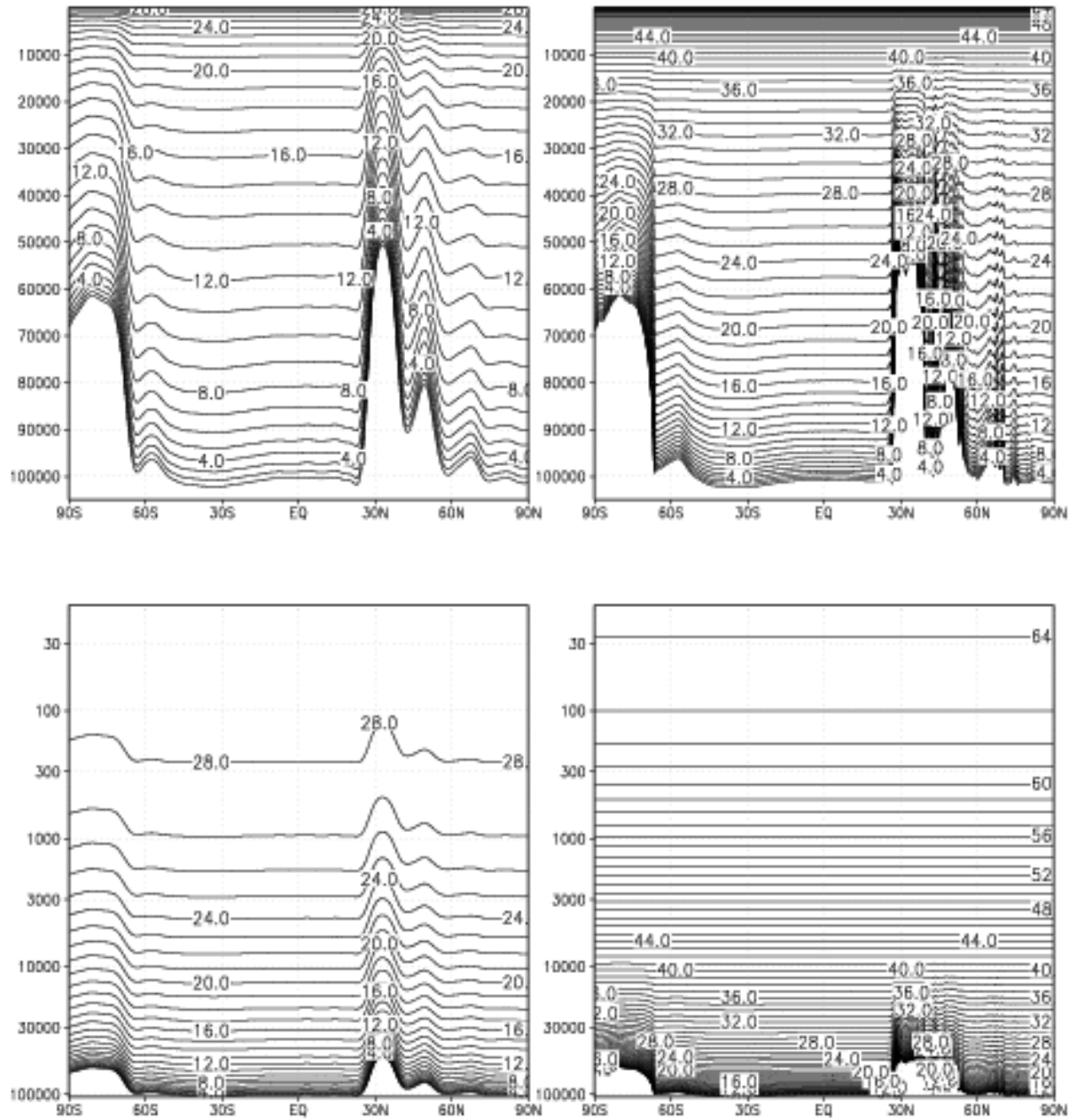


Figure 30: The vertical structure of model levels as a meridional cross section at 90E. The left panels are for R1 (28 sigma layers) and right panels for CFSR (64 sigma-pressure hybrid layers). The top panels are plotted as a linear function of pressure to emphasize resolution in the troposphere and bottom panels are plotted in log(pressure) to emphasize stratosphere.

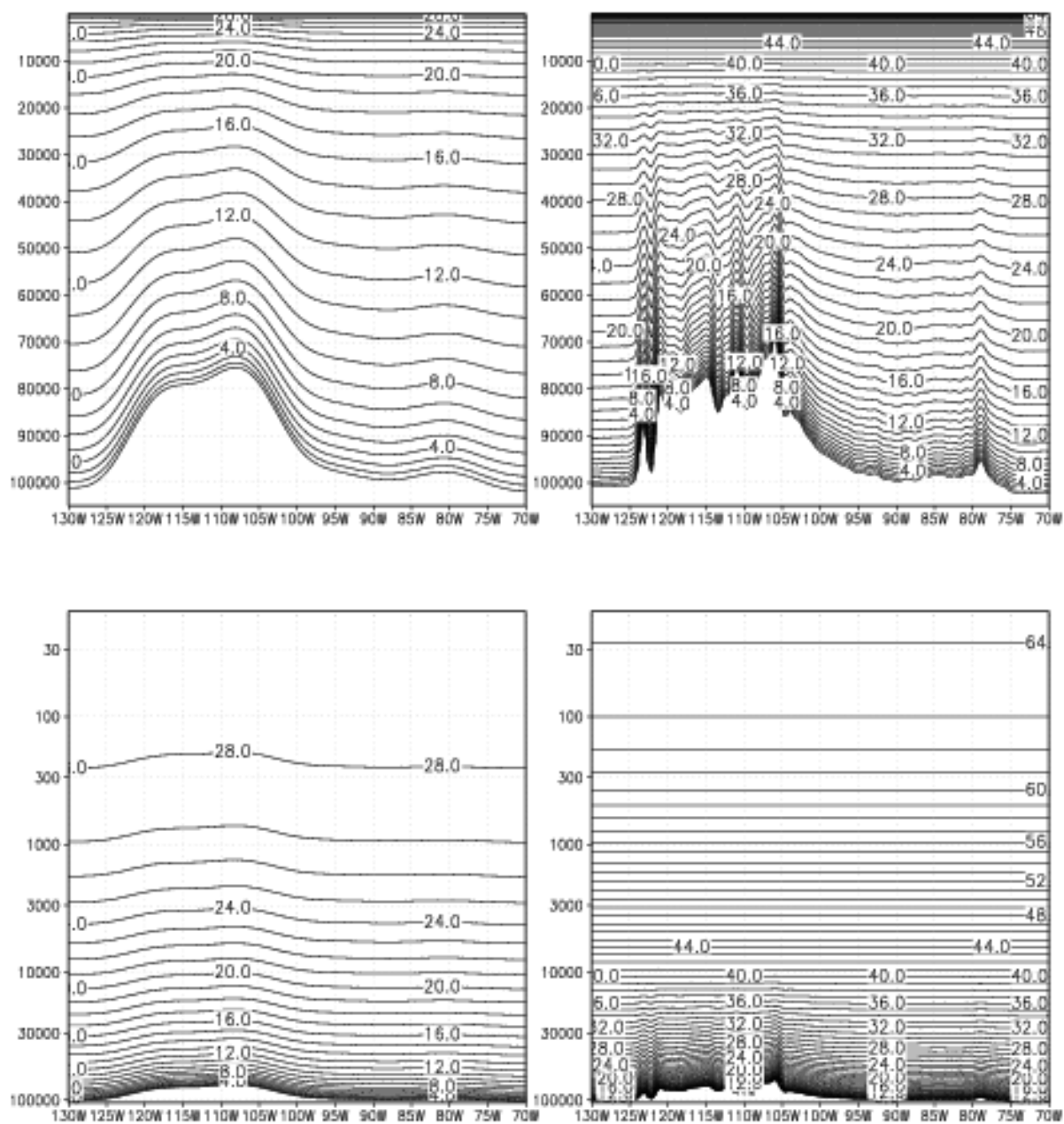


Figure 31: Same as in Figure 30, but as a zonal cross section across Rockies at 40N



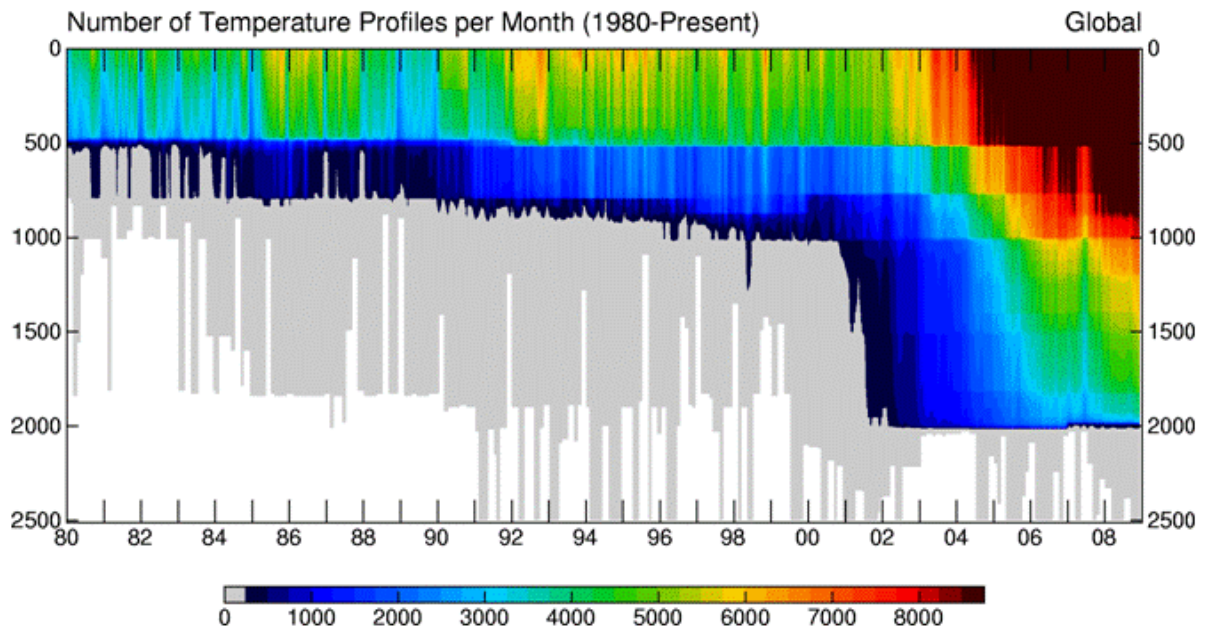


Figure 32: The global number of temperature observations assimilated per month by the ocean component of the CFSR as a function of depth for the years 1980 through 2008. The contour interval is 250 observations.



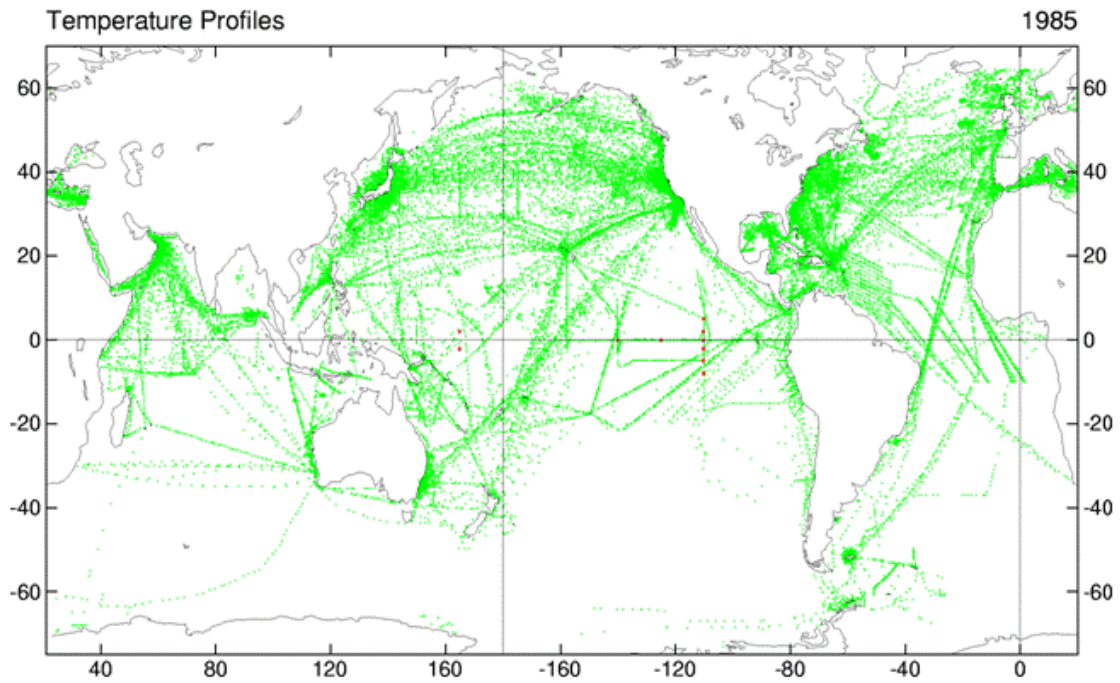


Figure 33: The global distribution of all temperature profiles assimilated by the ocean component of the CFSR for the year 1985

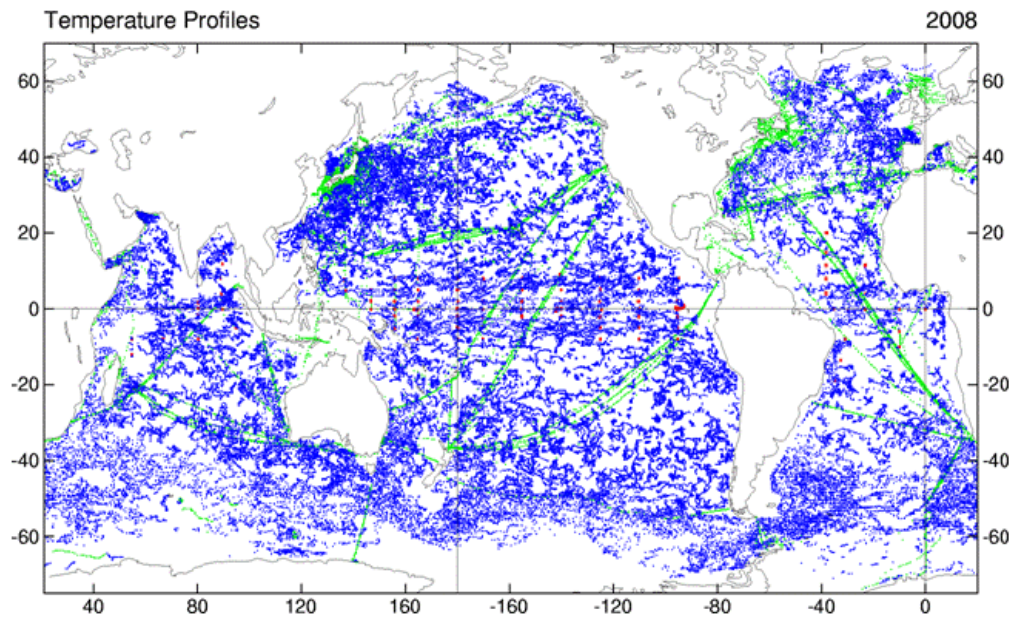


Figure 34: The same as Figure 33, but for the year 2008.

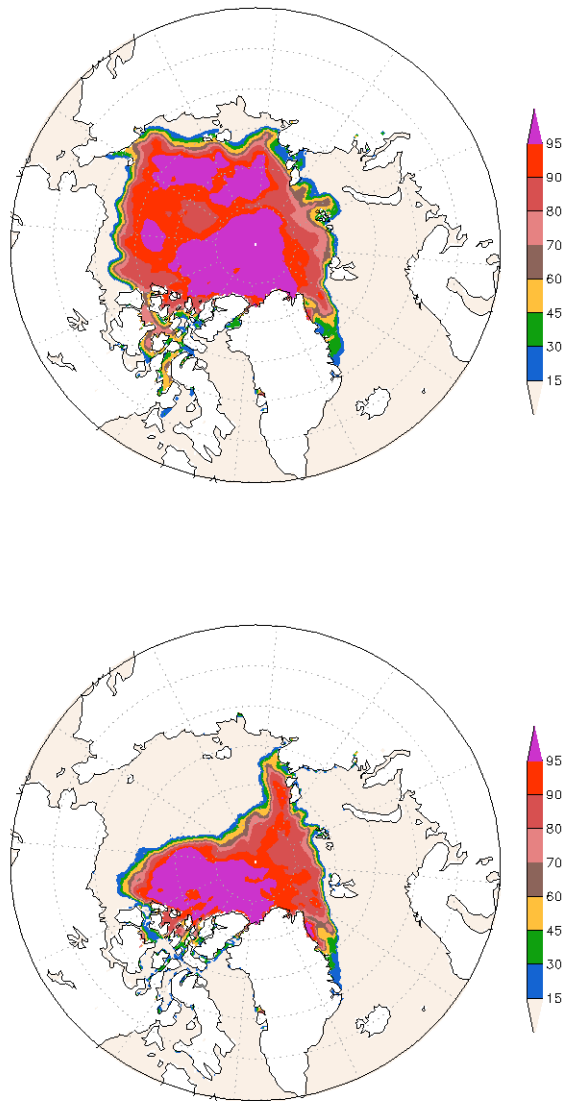


Figure 35: Sea ice concentration for September of 1987 (top) and 2007 (bottom) for the Arctic.

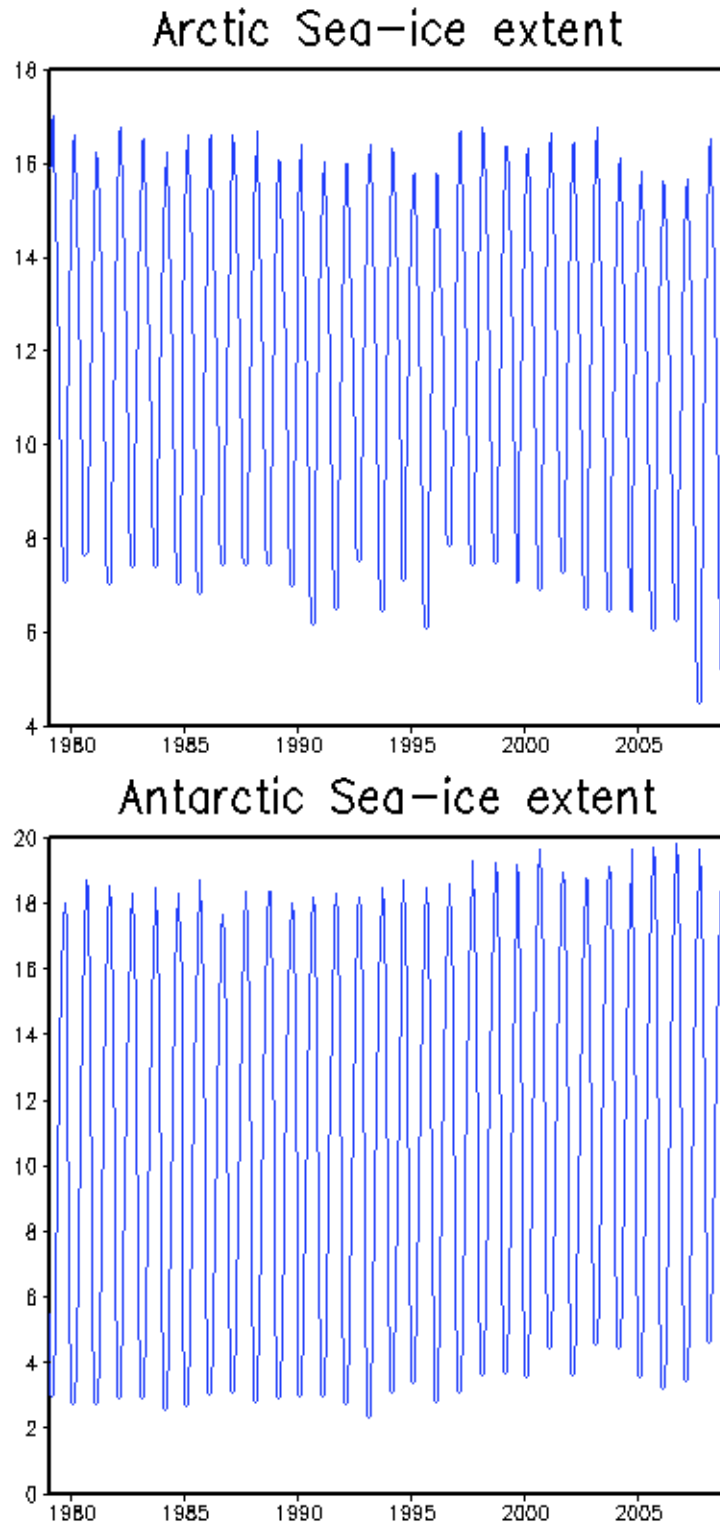


Figure 36: Monthly mean Sea ice extent ( $10^6 \text{ km}^2$ ) for the Arctic (top) and the Antarctic (bottom).

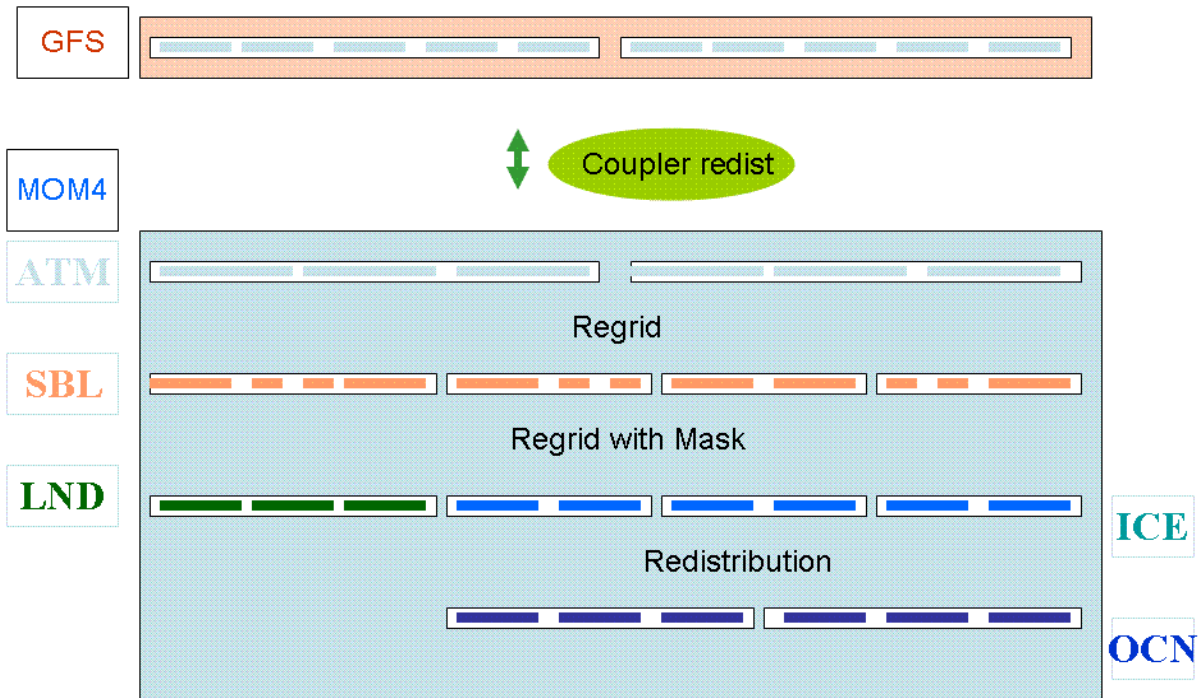


Figure 37: CFS grid architecture in the coupler. ATM is MOM4 atmospheric model (dummy for CFS), SBL is the surface boundary layer where the exchange grid is located, LAND is MOM4 land model (dummy for CFS), ICE is MOM4 sea ice model and OCN is MOM4 ocean model.

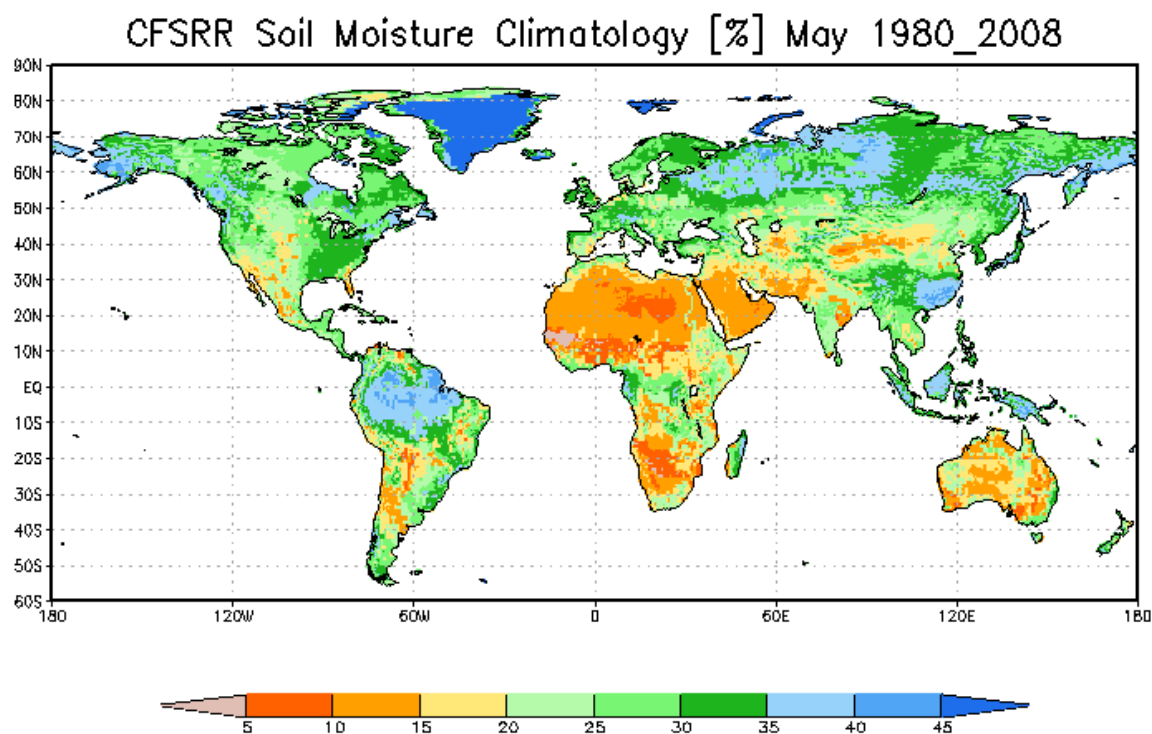


Figure 38: 2-meter volumetric soil moisture climatology of CFSR for May averaged over 1980-2008.

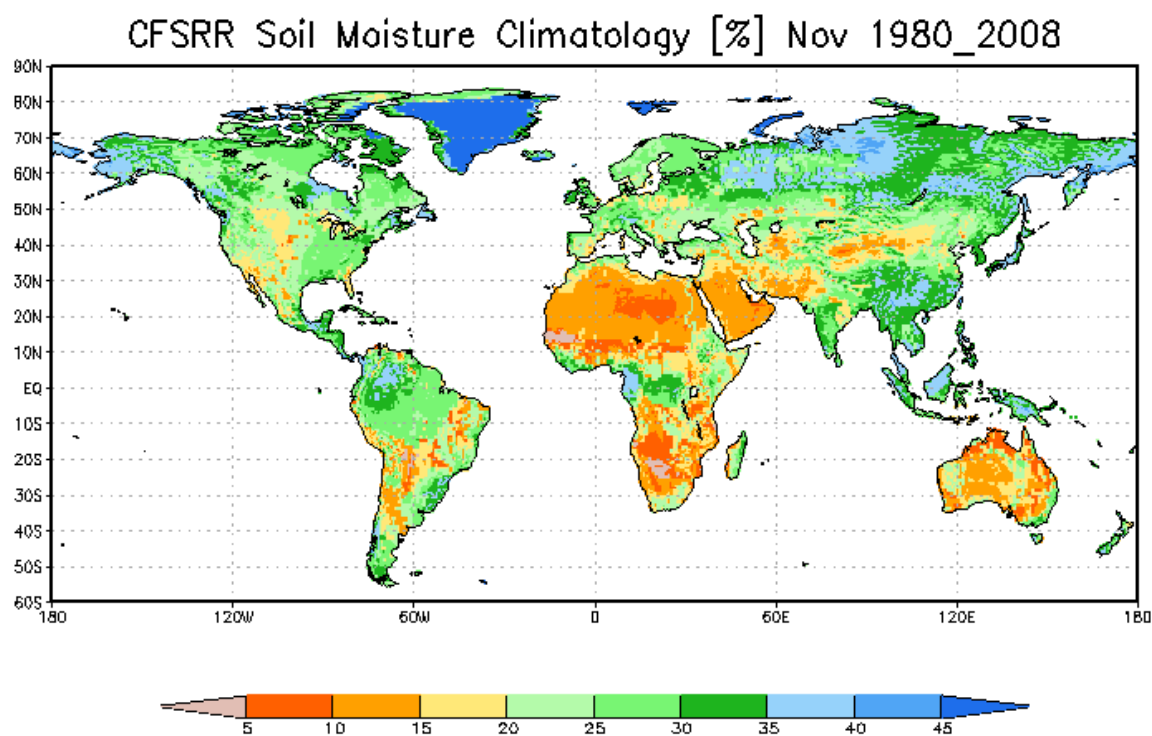


Figure 39: Same as in Figure 38, except for November.

## ONE DAY OF REANALYSIS

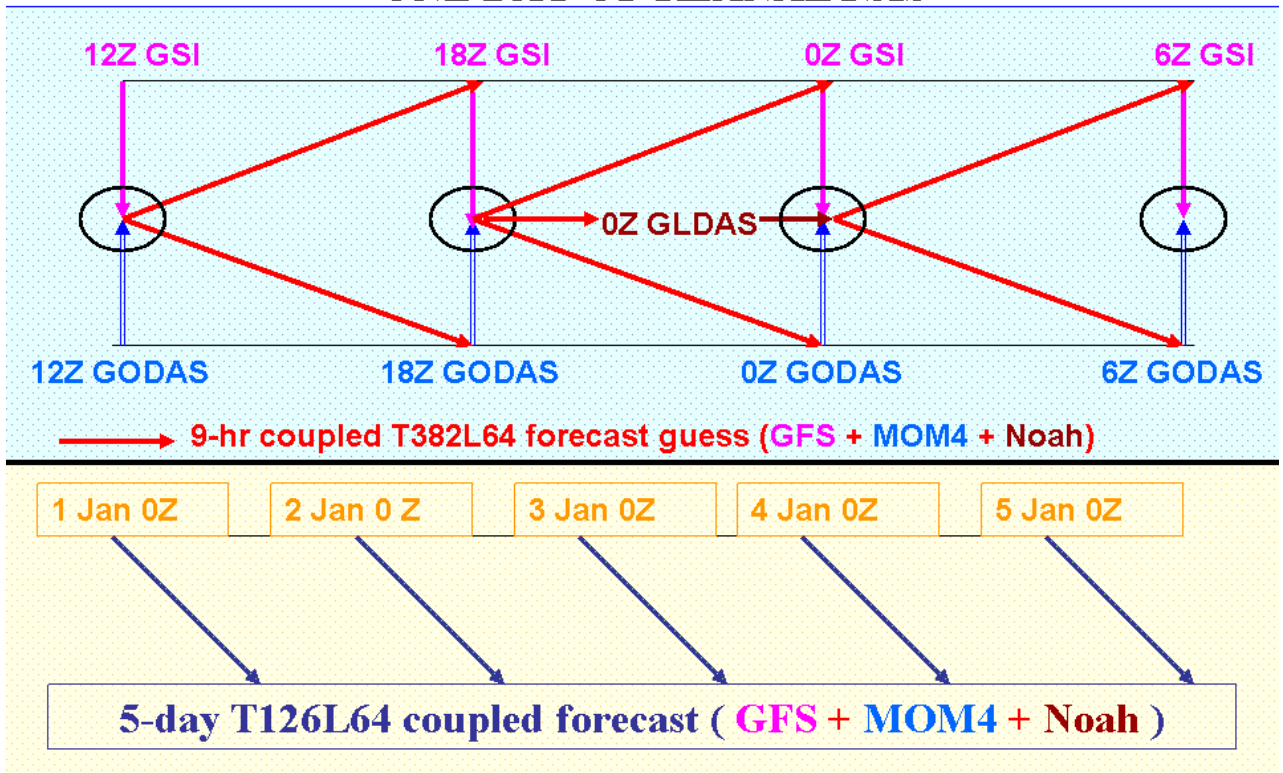


Figure 40: Schematic of the execution of one day of the CFS Reanalysis.



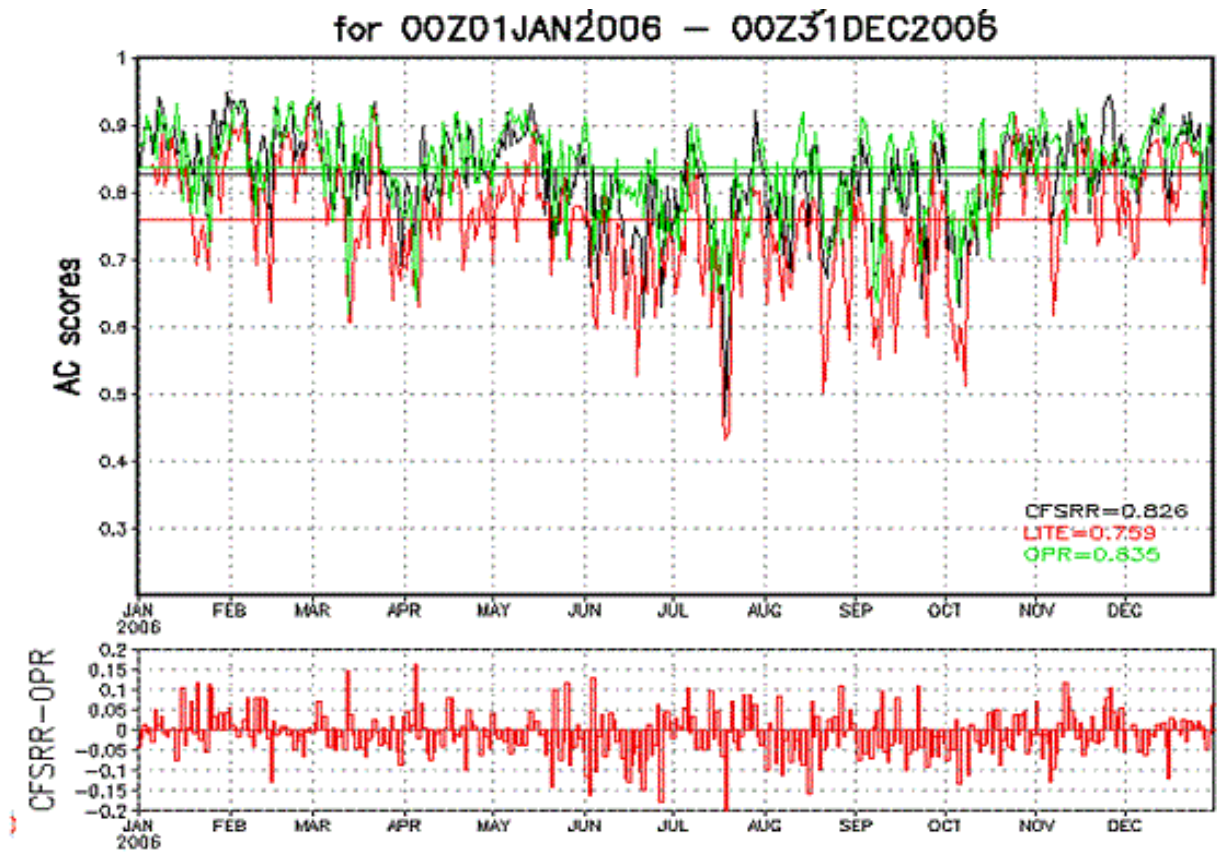


Figure 41: In the upper diagram are the daily time series of 0000 GMT 120 hour forecasts for the T126L64 CFSR (black), the T62L64 CFSR-Lite (red) , and operational (OPR) T382L64 operational GFS for 2006. The lower diagram is the time series of the difference CFSR minus GFS.

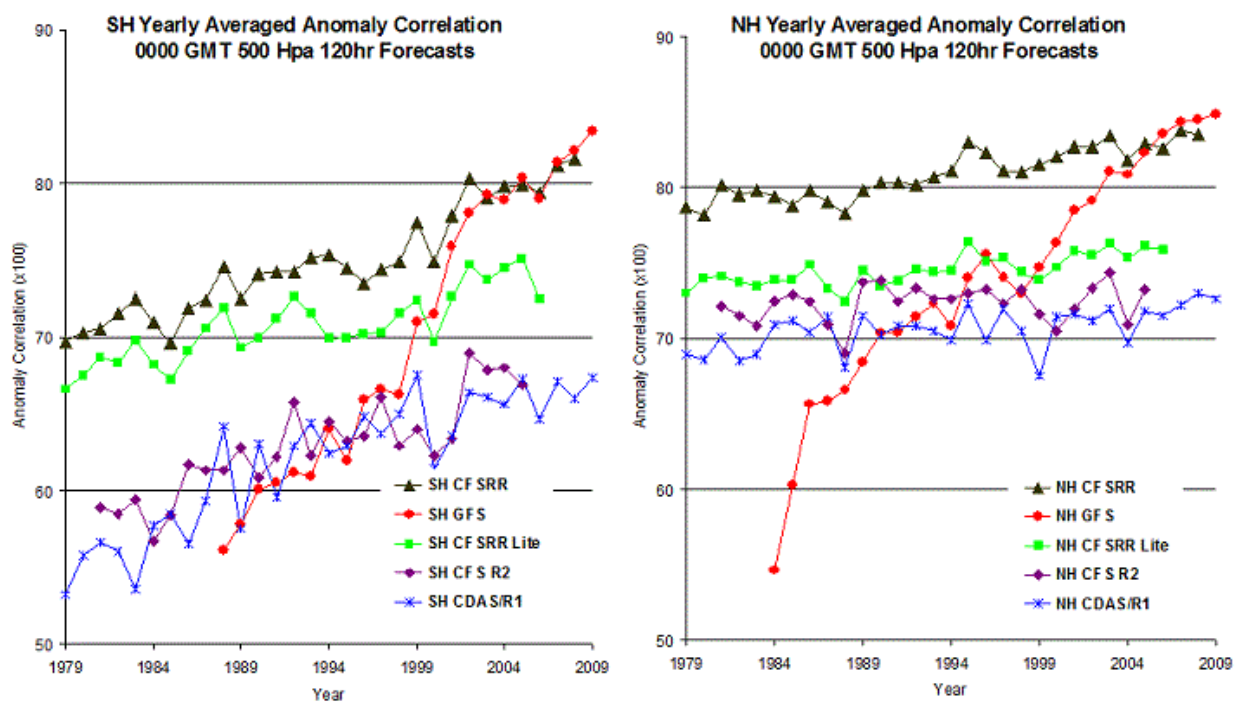


Figure 42: Yearly averaged Southern Hemisphere (left) and Northern Hemisphere (right)

00 GMT 120 hour forecast anomaly correlations for CFSR (black triangles),  
GFS (red circles), CFSR-Lite (green squares), CFS R2 (purple diamonds),  
CDAS R1 (blue stars).

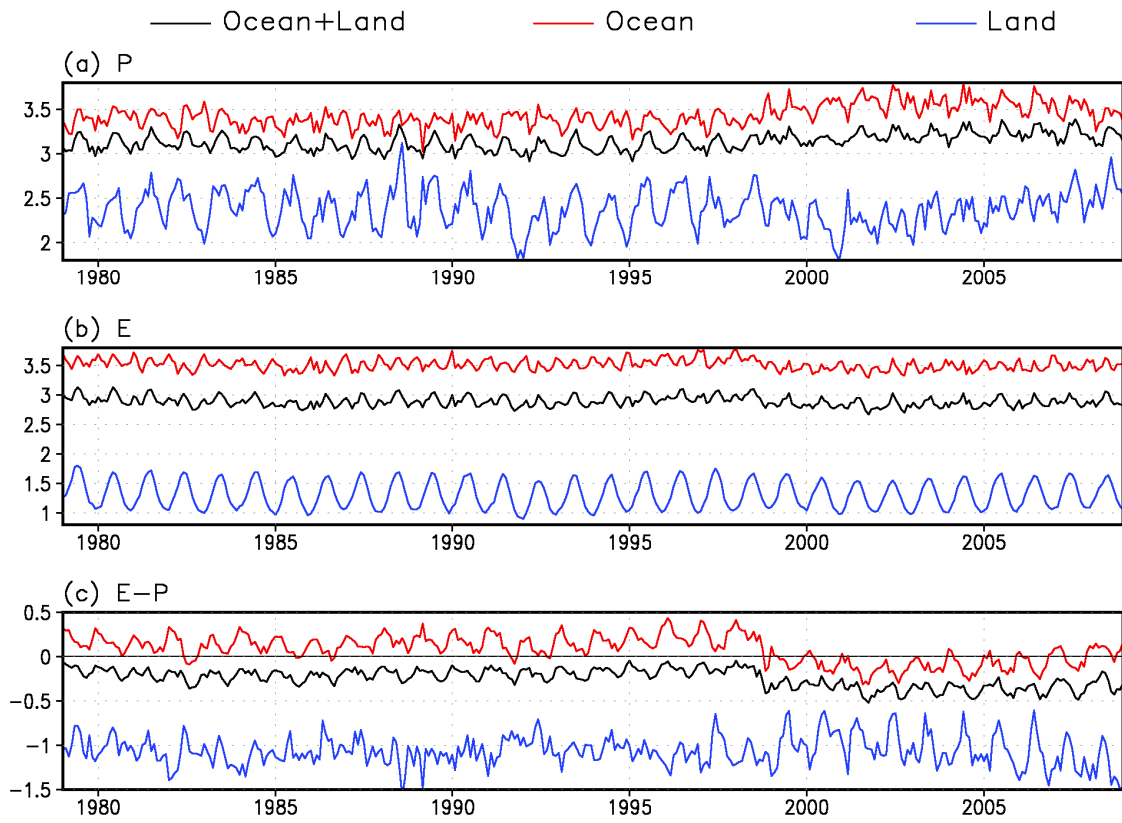


Figure 43: Global average of monthly-mean precipitation (a), evaporation (b) and evaporation minus precipitation (c). Averages over ocean, land, and ocean plus land are plotted with red, blue, and black curves. Unit is mm/day.

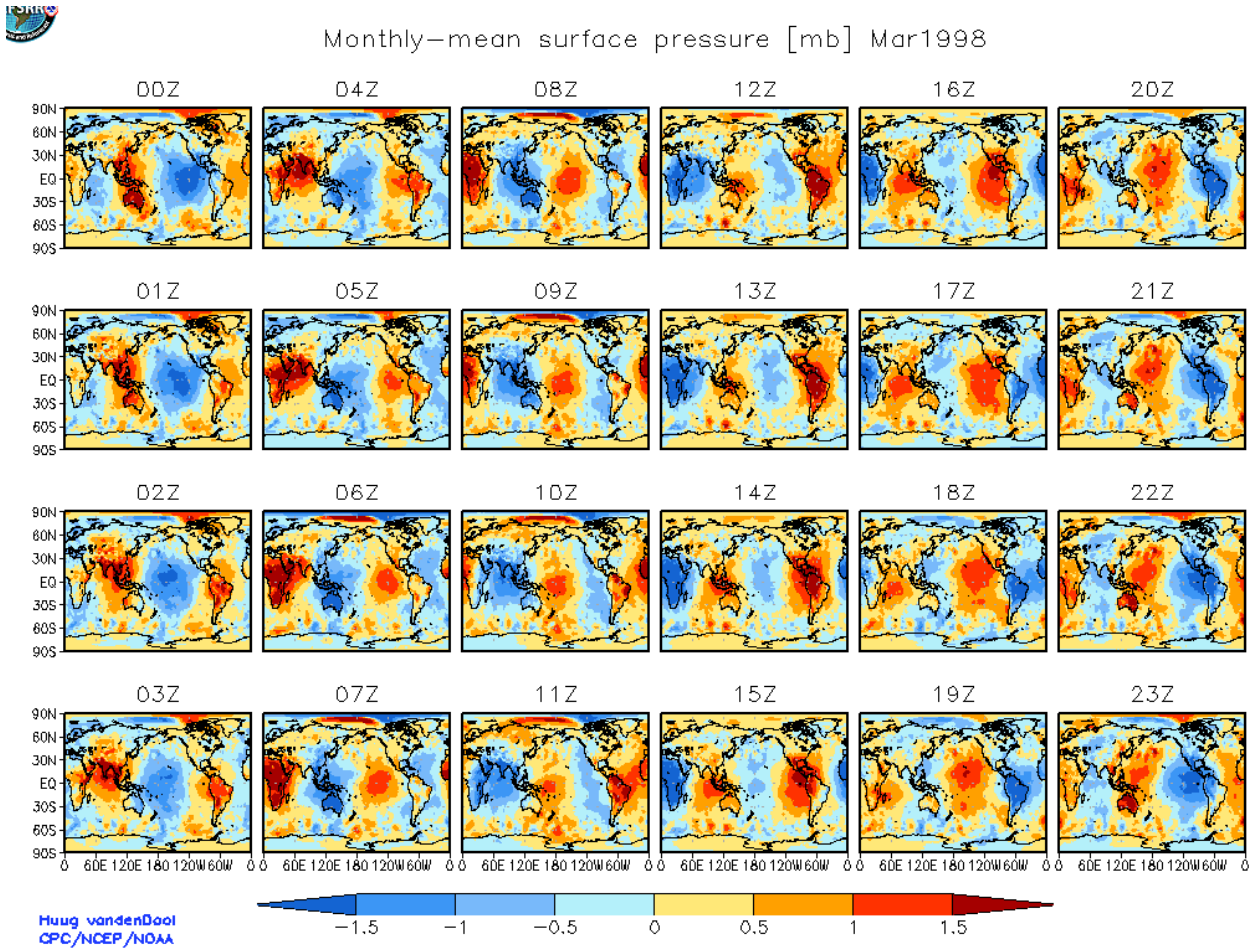


Figure 44: Monthly mean hourly surface pressure (with the daily mean subtracted out) for March 1998 in a collage of 24 global maps. Each map is on the T382 Gaussian grid. In red (blue) areas pressure is higher (lower) than the daily mean. Units are hPa. Time starts in the upper left (00Z), then proceeds down to 3Z, then continues at the top of the next column of global maps etc.

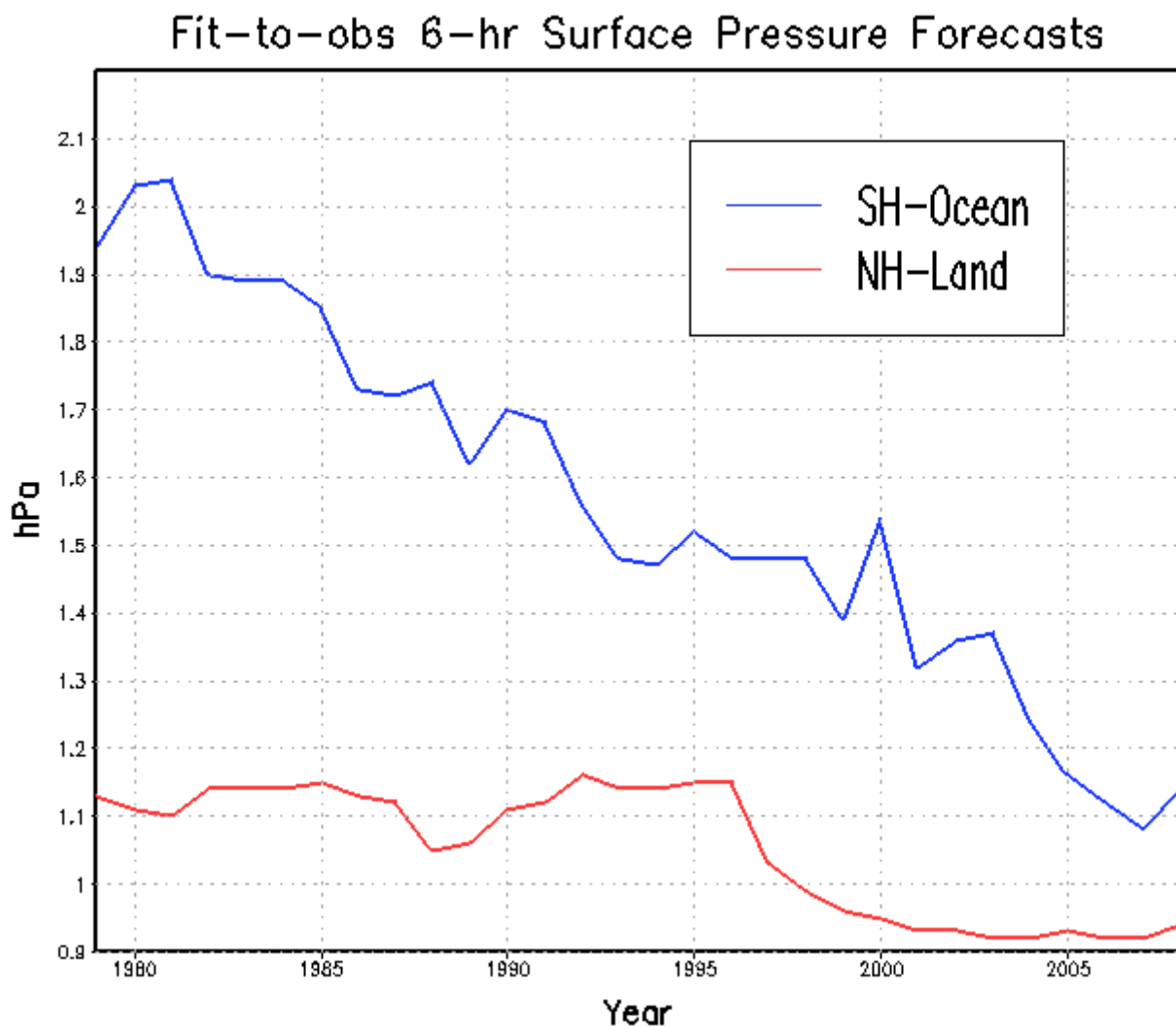


Figure 45: The fit of 6-hour forecasts of instantaneous surface pressure against irregularly distributed observations also called the background error. Shown are annually compiled fits-to-obs data 1979-2008 in units of mbar. In blue SH ocean, in red NH land.

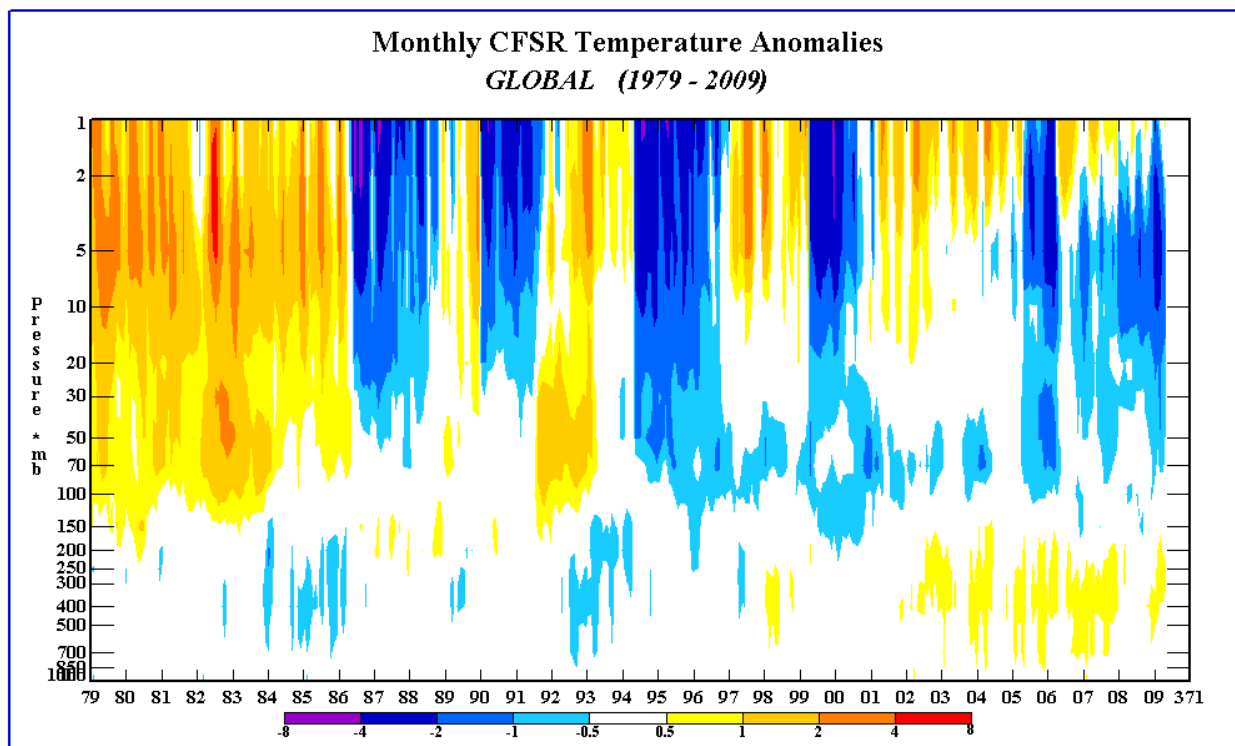


Figure 46: Global mean temperature anomalies from 1000 to 1 hPa from January 1979 through May 2009. Units are in K.

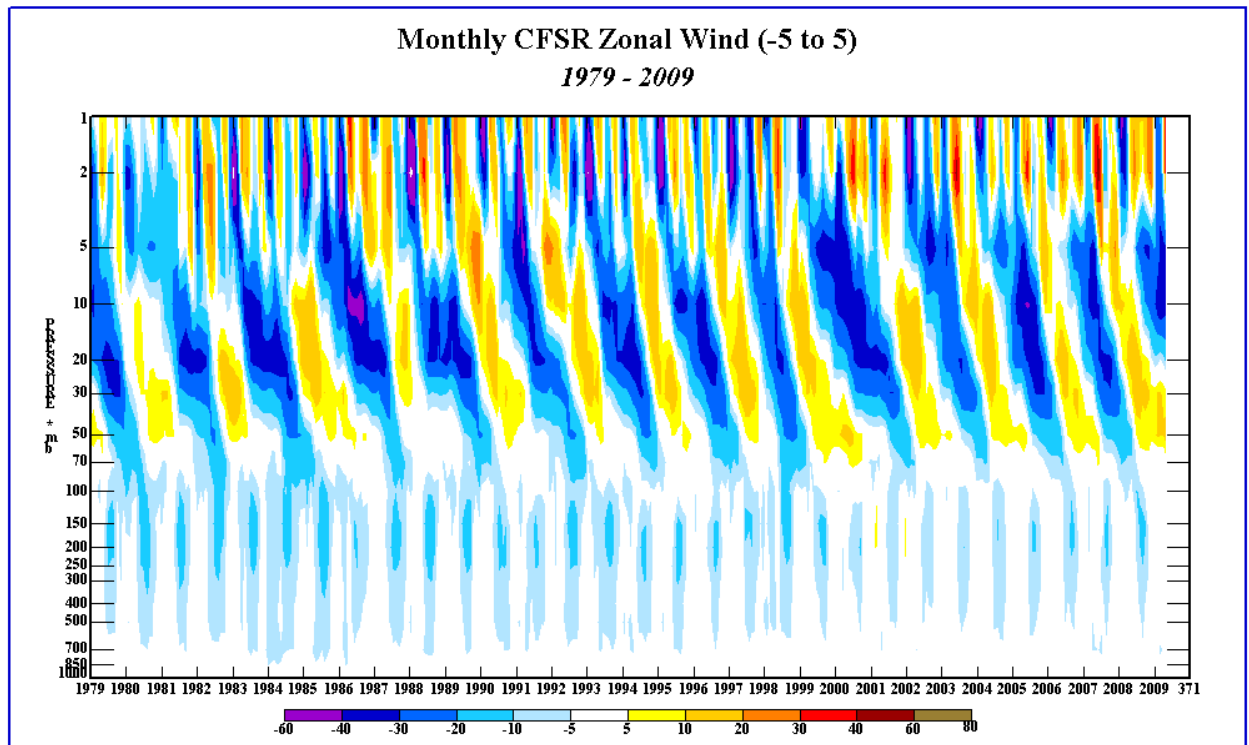


Figure 47: Equatorial (5°S to 5°N) zonal mean U wind component time series from 1979 to 2009. Units are in m/sec.

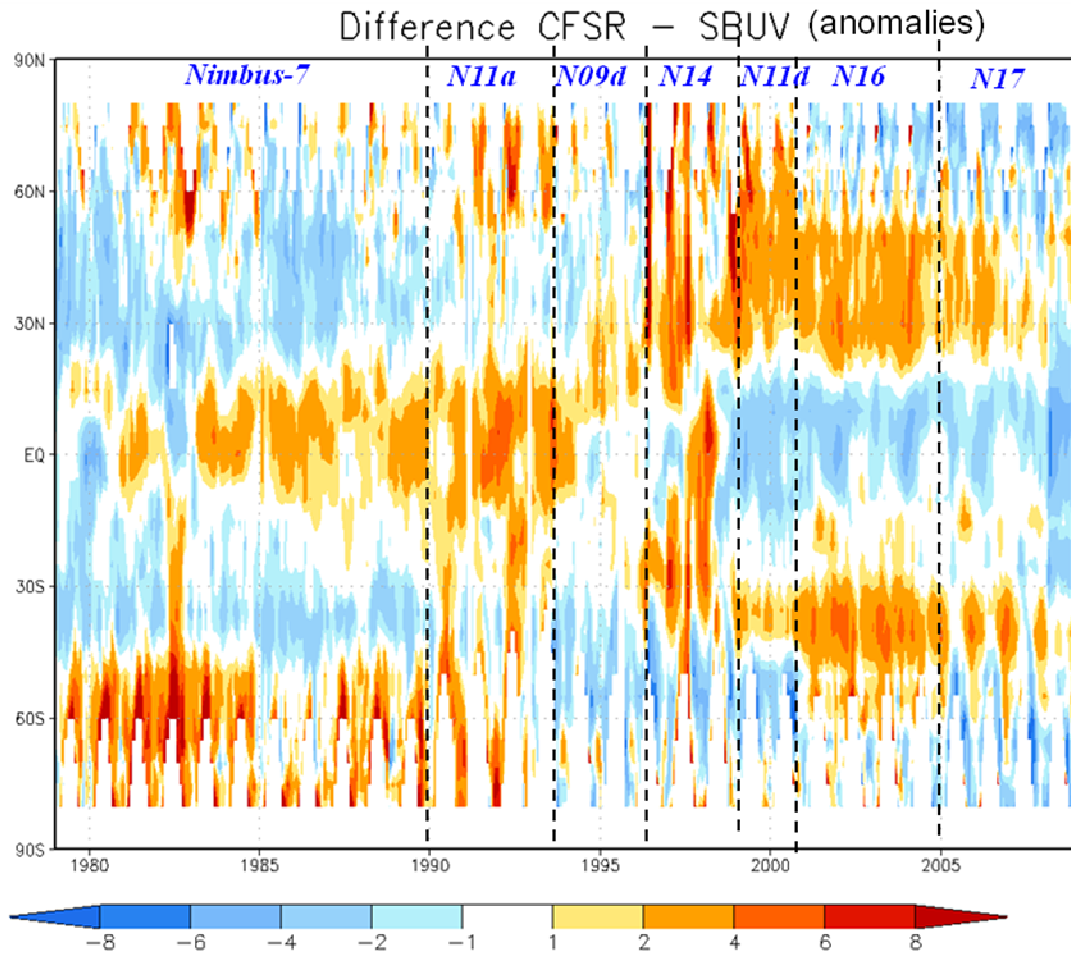


Figure 48: Zonal mean total ozone differences between the CFSR and the Cohesive SBUV(/2) Total Ozone Dataset. Units are in Dobson units. Also shown are the time spans of each satellite's SBUV(/2) data used in the Cohesive SBUV(/2) Total Ozone Dataset.



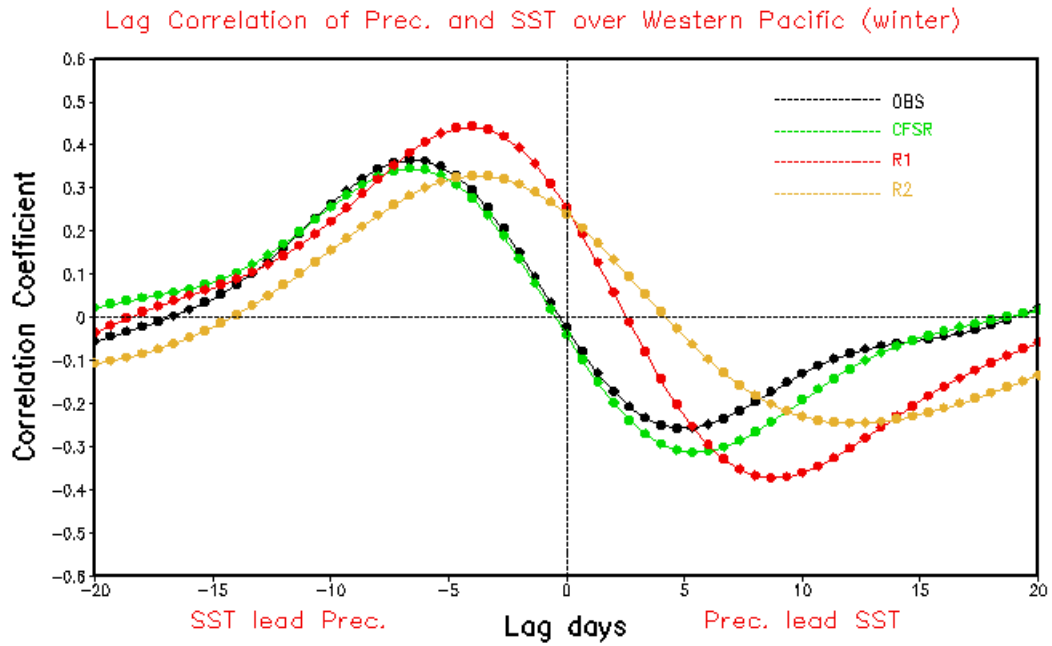


Figure 49: Temporal lag correlation coefficient between precipitation and SST in the tropical Western Pacific (averaged over 10S-10N, 130E-150E) in R1 (red), R2 (brown), CFSR (green) and observation (black). GPCP daily precipitation and Reynolds  $\frac{1}{4}^\circ$  daily SST are used as observational data. Negative (positive) lag days on the x-axis indicate the SST leads (lags) the precipitation. Data for the boreal winter (Nov-Apr) over the period 1979-2008 are band-pass filtered for 20-100 days after removing the climatological mean.

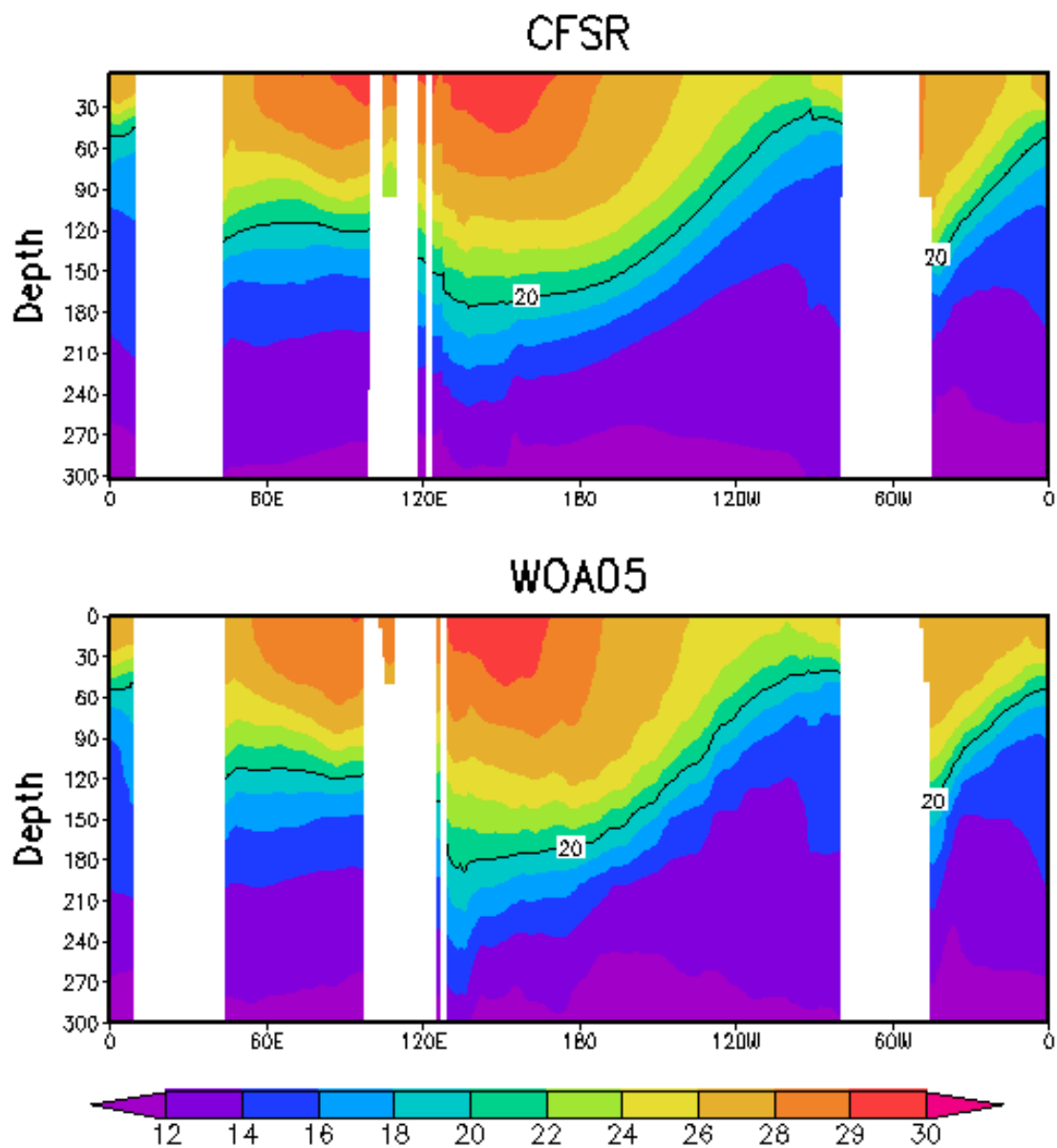


Figure 50: The subsurface temperature mean for an equatorial cross-section for CFSR (top) and observations from the World Ocean Atlas (Conkright et al., 1999) .

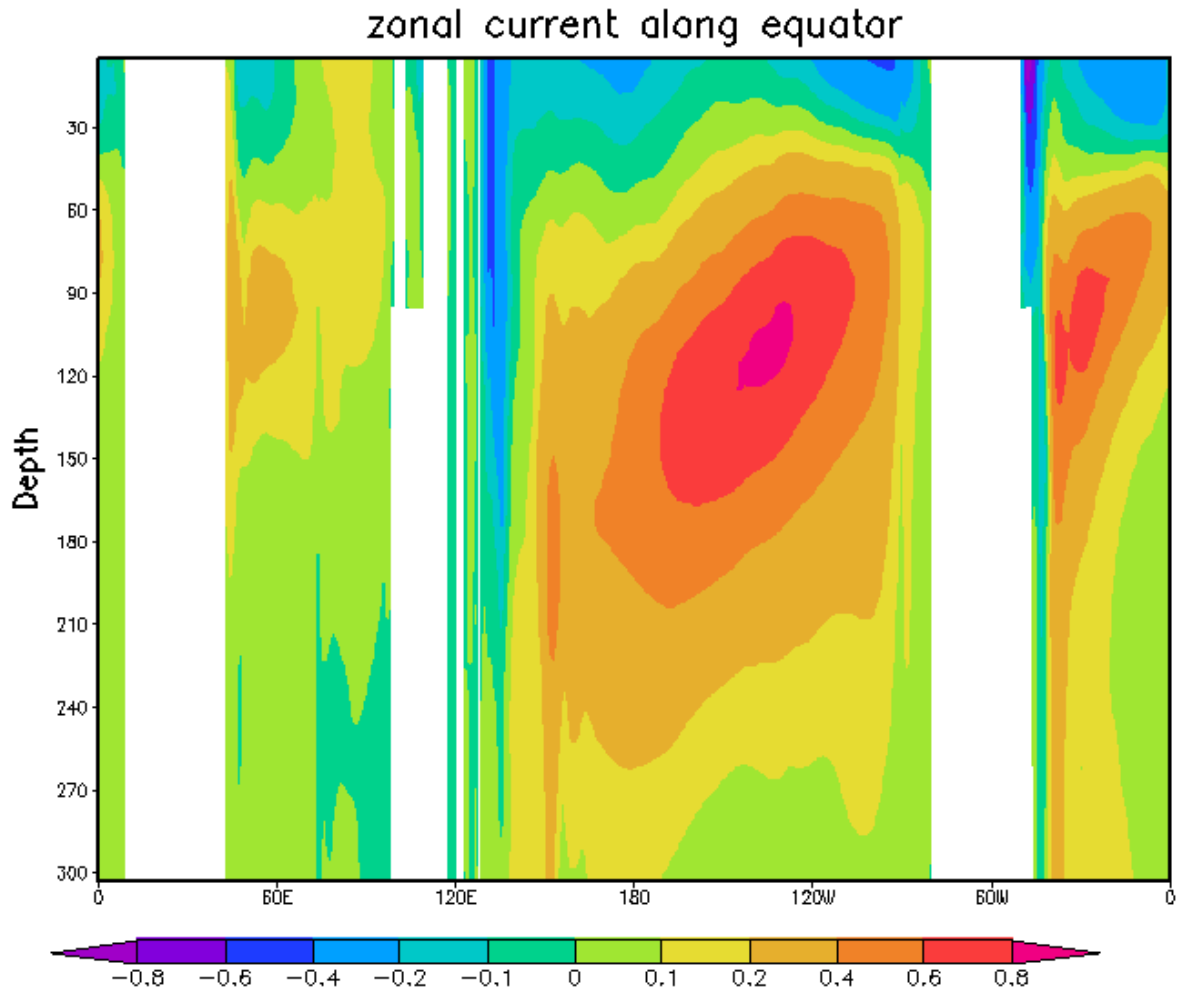


Figure 51: The subsurface zonal velocity for an equatorial cross-section for CFSR for the years: 1979-2008.

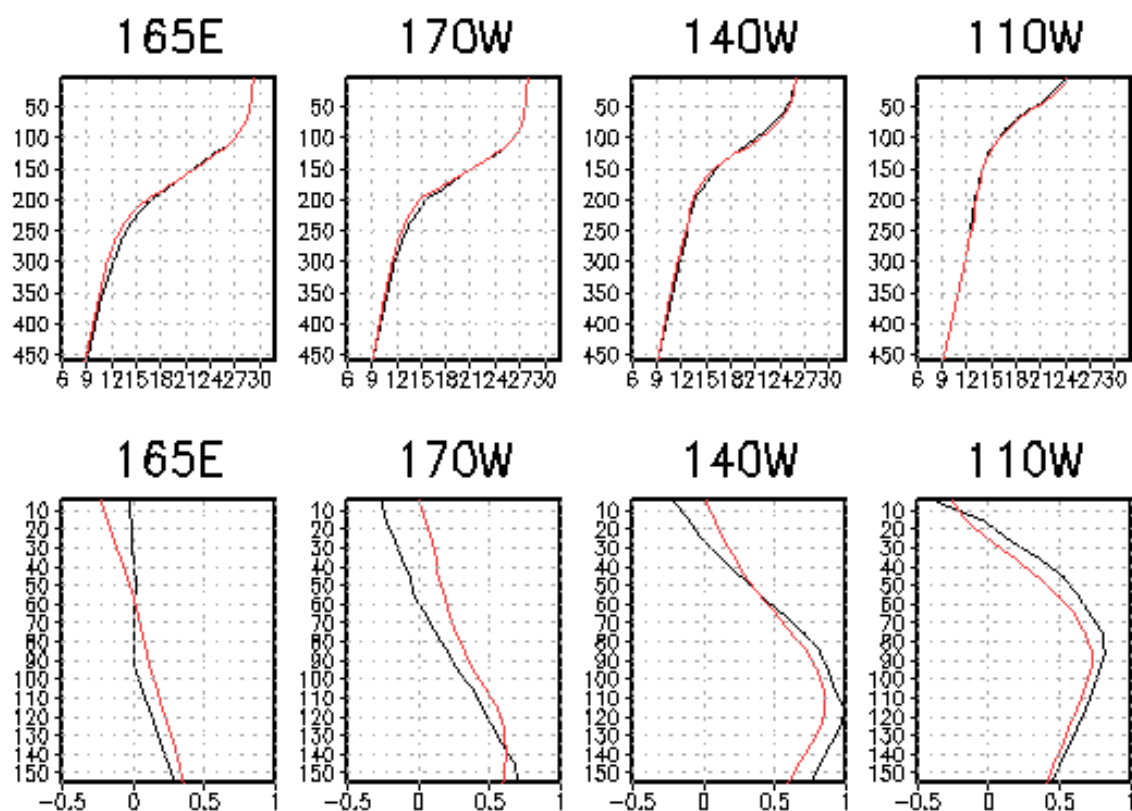


Figure 52: Vertical profiles of the subsurface temperature (top panels) and zonal velocity (bottom panels) for CFSR (red line) and TAO observations (black line) for four locations in the equatorial Pacific ocean.

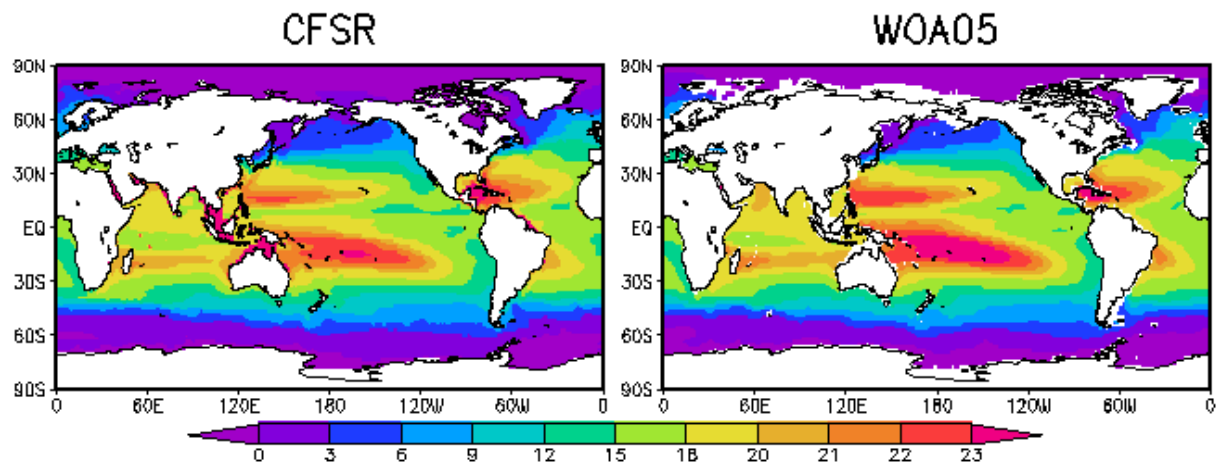


Figure 53: The vertically averaged temperature (top 300 m) for CFSR for 1979-2008, and for observations from the World Ocean Atlas (Conkright, et al., 1999).

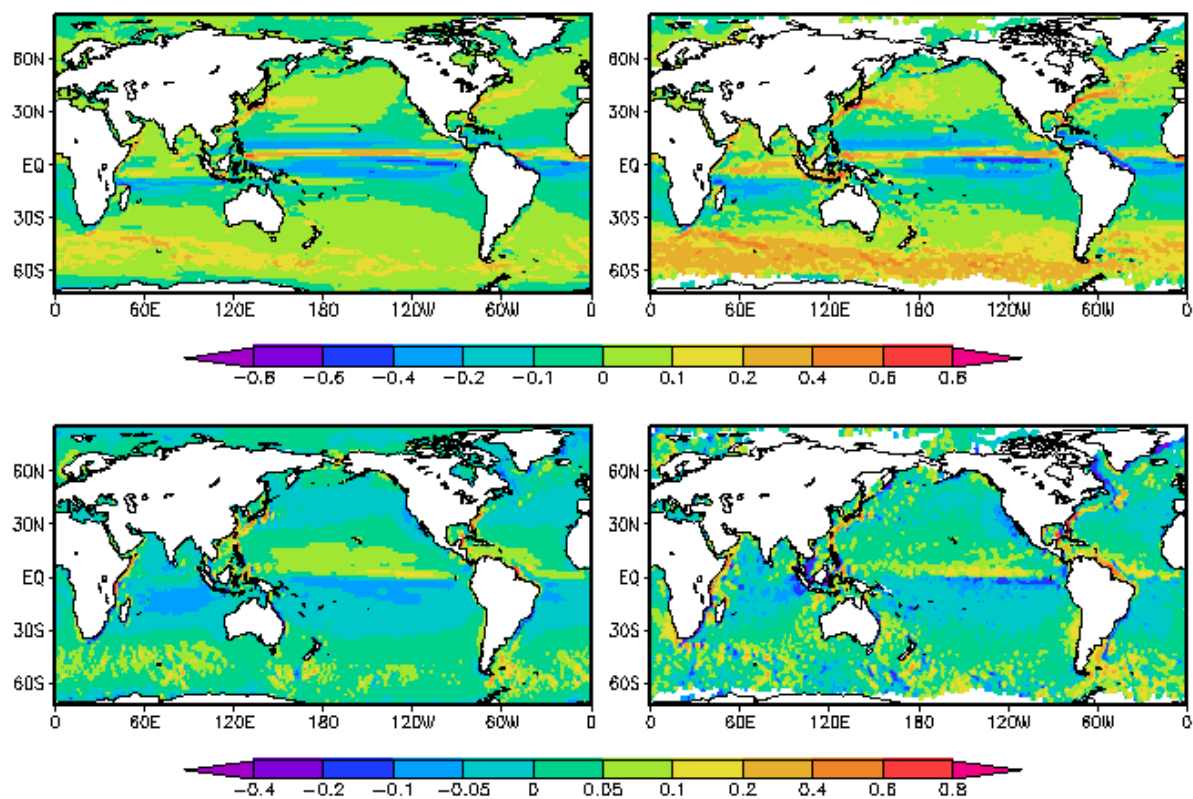


Figure 54: Zonal surface velocities for CFSR ( top left) and for drifters from the Surface Velocity Program of TOGA (top right). Meridional velocities are shown in the bottom panel.

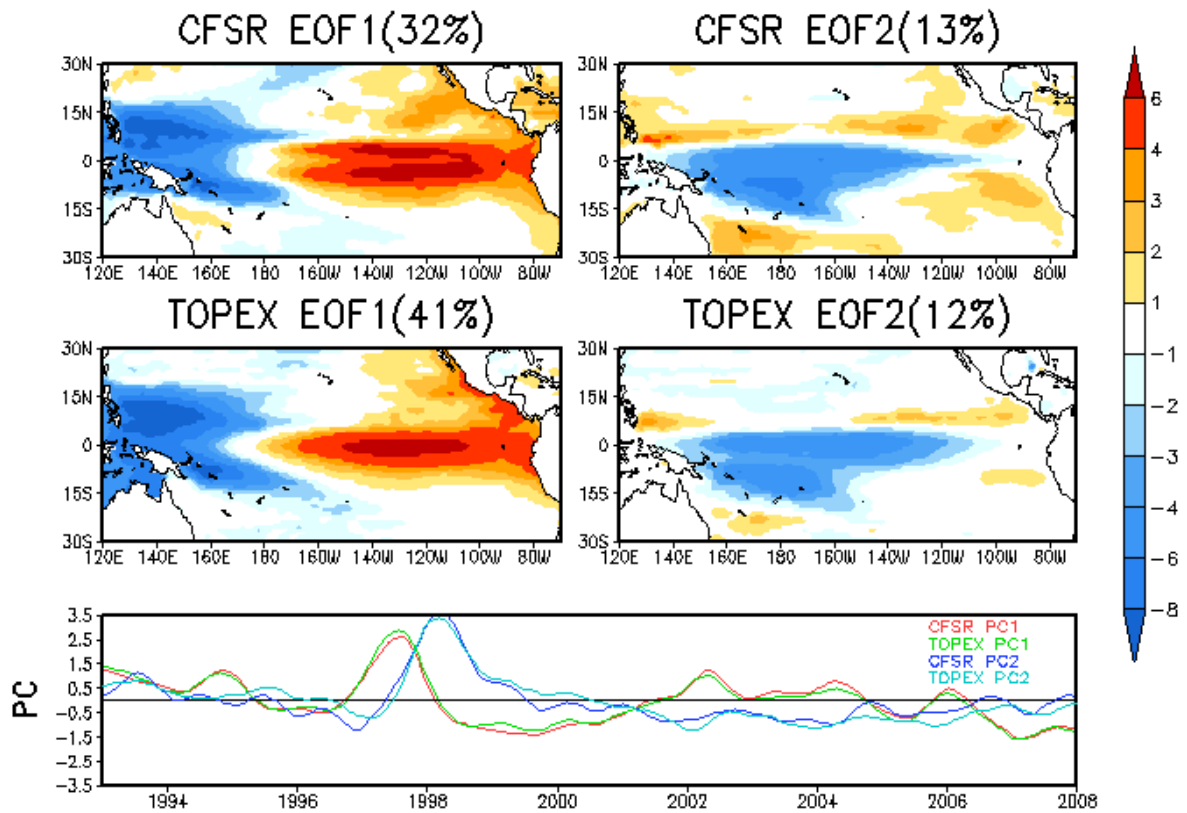


Figure 55: The first two EOFs of the SSH variability for the CFSR (left) and for TOPEX satellite altimeter data (right) for the period: 1993-2008. The time series amplitude factors are plotted in the bottom panel.

Program	Purpose
PREVENTS	Store first guess values as meta data Store observational errors as metadata Surface pressure data gross checks Observation screening via screening list
ACQC	Aircraft quality control, track check, and duplicate check
ACAR_CQC	ACARS aircraft ascent and descent check
CQC	Radiosonde complex quality control (Gandin, 1988, Collins and Gandin, 1990) Calculate radiation bias correction
CQCVAD	NEXRAD VAD wind complex quality control
PROFCQC	Wind profiler complex quality control
GSI	Non-linear variational observation quality control (VarQC)
POSTEVENT	Store VarQC results as metadata Store analysis values as metadata

Table 1: NCEP operational observation quality control procedures



<b>Satellite Platform</b>	<b>Launch Date<sup>a</sup></b>	<b>Mission End Date<sup>a</sup></b>	<b>Mission Mean (ppmv)<sup>b</sup></b>
TIROS-N	13 Oct 78	1 Nov 80	337.10
NOAA-6	27 Jun 79	19 Sep 83	340.02
NOAA-7	23 Jun 81	7 Feb 85	342.86
NOAA-8	28 Mar 83	26 May 84	343.67
NOAA-9	24 Dec 84	13 Feb 98	355.01
NOAA-10	17 Sep 86	17 Sep 91	351.99
NOAA-11	24 Sep 88	Jun 04	363.03
NOAA-12	14 May 91	Standby	365.15
GOES-8	13 Apr 94	5 May 04	367.54
GOES-9	23 May 95	28 Jul 98	362.90
GOES-10	25 Apr 97	West (in transition)	370.27
NOAA-15 to NOAA-18	26 Oct 98	20 May 05	380.00
IASI (Metop-A)	19 Oct 06	Current	389.00

Table 2: Carbon Dioxide Concentrations (ppmv) used in the CRTM to calibrate satellites.

<sup>a</sup> Davis (2007)

<sup>b</sup> Historical CO<sub>2</sub> observed data available at WMO Global Atmospheric Watch

<http://gaw.kishou.go.jp>

Satellite	Starting date	Ending Date
<b>N6-1</b>	<b>1979070100</b>	<b>1979100100</b>
<b>TIROS-N</b>	<b>1979010100</b>	<b>1979040100</b>
<b>N7</b>	<b>1981080100</b>	<b>1981120100</b>
<b>N8</b>	<b>1983050100</b>	<b>1983080100</b>
<b>N9</b>	<b>1985022500</b>	<b>1985060100</b>
<b>N6-2</b>	<b>1985101800</b>	<b>1986020100</b>
<b>N10</b>	<b>1986112600</b>	<b>1987030100</b>
<b>N11</b>	<b>1988110800</b>	<b>1989020100</b>
<b>N12</b>	<b>1991091600</b>	<b>1992010100</b>
<b>N14</b>	<b>1995010300</b>	<b>1995080100</b>
<b>N11-2, G08, G09</b>	<b>1997071500</b>	<b>1998010100</b>
<b>N15,G10</b>	<b>1998091000</b>	<b>1999020100</b>
<b>AMSub_N15</b>	<b>2000010600</b>	<b>2000040100</b>
<b>N16</b>	<b>2001012400</b>	<b>2001060100</b>
<b>G12</b>	<b>2003040200</b>	<b>2003070100</b>
<b>G11</b>	<b>2006062100</b>	<b>2006100100</b>

Table 3: List of experiments for the spin up of bias correction of historical satellites with their corresponding starting and ending times (N is for NOAA and G is for GOES in the first column).